UNIVERSITÀ DI PISA

FACOLTÀ DI INGEGNERIA

RELAZIONE PER IL CONSEGUIMENTO DELLA

LAUREA SPECIALISTICA IN INGEGNERIA GESTIONALE

"From the conception to the definition of a new service: the case of the European GeoPKDD project"

RELATORI

IL CANDIDATO

Prof. Giuseppe Bellandi

Dipartimento Sistemi Elettrici e Automazione

Matteo Piazza Bonati

Ing. Andrea Chiarello

Tutor Aziendale

Anno Accademico 2006-2007

Abstract

The thesis shows the process of generation and definition of a new service through the analysis of the work done for Wind Telecomunicazioni S.p.A in the context of the European project GeoPKDD (Geographic Privacy-aware Knowledge Discovery and Delivery). The contribute that this work aims to provide is a marketing oriented view inside the project that, for its nature, is technologically oriented, in order to show how the results obtained so far can be embodied in a service that meets the market's requirements. Starting from the examination of the technology at disposal, which aims to analyse users' movement through data coming from their mobile devices, the applicative part of the thesis is divided in two main phases. In the first one, that is service generation, the work illustrates the process that brought to the creation of several innovative service ideas and that, through a screening and a scoring phases, ended with the selection of the most promising idea to develop. Then, in the service definition phase, target market is analysed and, through the gathering of needs from some potential clients that were directly contacted, the service is defined in terms of functionalities to implement and value configuration for the final client.

Sommario

La tesi illustra il processo di generazione e di definizione di un nuovo servizio attraverso l'analisi del lavoro svolto per WIND Telecomunicazioni S.p.A. nell'ambito del progetto Europeo GeoPKDD (Geographic Privacy-aware Knowledge Discovery and Delivery). Il contributo che questo lavoro si propone di fornire è una visione orientata al marketing all'interno del progetto che, per sua natura, è incentrato su aspetti tecnologici, al fine di mostrare come i risultati fin qui ottenuti possano essere incorporati in un servizio che incontri le richieste del mercato. Partendo dall'esame della tecnologia a disposizione, il cui fine è l'analisi del movimento degli utenti attraverso dati provenienti dai loro dispositivi mobili, il lavoro è suddiviso in due macrofasi. Nella prima, che è la generazione del servizio, viene illustrato il processo che ha portato alla creazione di numerose idee innovative e che, attraverso una fase di screening e poi di scoring, è terminato con la selezione dell'idea più promettente da sviluppare. In seguito, nella fase di definizione, il mercato di riferimento viene analizzato e, attraverso la raccolta dei bisogni di potenziali clienti direttamente contattati, il servizio viene definito in termini di funzionalità da implementare e di configurazione del valore per il cliente finale.

Table of Content

SOMMARIO 2

INTRODUCTION AND STRUCTURE OF THE WORK 6

1 NEW TECHNOLOGY PUSH DEVELOPMENT MODELS 8

1.1 NEW SERVICE DEVELOPMENT 8
1.2 NEW TECHNOLOGY-PUSH SERVICES 12
1.3 TECHNOLOGY PUSH DEVELOPMENT MODELS 16
1.4 STRATEGY IN TECHNOLOGY PUSH MODELS 20

2 NEW SERVICE CONCEPTION AND DEFINITION 28

2.1 GENERATION 29
2.2 CONCEPT SCREENING 40
2.3 CONCEPT SCORING 42
2.4 SERVICE DEFINITION 45
2.4.1 INITIAL MARKET ANALYSIS 45
2.4.2 GATHER INFORMATION FROM THE CLIENT 52
2.4.3 IDENTIFY AND ORGANIZE NEEDS 55
2.4.4 DEFINE PRODUCT FUNCTIONALITIES 57
2.4.5 SELECT THE FUNCTIONALITIES TO IMPLEMENT 58

3 OVERALL ON THE EUROPEAN GEOPKDD PROJECT 62

3.1 OBJECTIVES OF THE PROJECT 62
3.2 THE GEOPKDD PROCESS 64
3.2.1 PRIVACY-AWARE TRAJECTORY WAREHOUSE 65
3.2.2 PRIVACY-AWARE SPATIO TEMPORAL DATA MINING METHODS 66
3.2.3 GEOGRAPHIC KNOWLEDGE INTERPRETATION AND DELIVERY 67
3.2.4 APPLICATION DEMONSTRATORS 68
3.2.5 HARMONIZATION AND DISSEMINATION 68

3.3 ANALYSIS OF GEOPKDD PROCESS 69
3.3.1 INPUT 70
3.3.2 INSIDE THE PROCESS 73
3.3.3 CONSTRAINTS 77
3.3.4 RESOURCES 81
3.3.5 OUTPUT 84
3.4 WHERE THIS WORK STANDS 90

4 CONCEPT GENERATION FOR THE GEOPKDD PROJECT 92

4.1 GENERATION OF IDEAS 92
4.2 CONCEPT SCREENING 98
4.3 CONCEPT SCORING AND SELECTION 106

5 DEFINITION OF FUNCTIONALITIES FOR THE GEOPKDD APPLICATION 110

5.1 INITIAL MARKET ANALYSIS 110
5.1.1 MARKET IDENTIFICATION 110
5.1.2 SWOT ANALYSIS 114
5.1.3 MARKET SEGMENTATION 116
5.1.4 DEMAND FORECAST 118
5.1.5 INDUSTRY IDENTIFICATION 119
5.2 GATHER INFORMATION FROM THE CLIENT 131
5.2.1 AMA, MOBILITY AGENCY OF MILAN 134
5.2.2 COMUNE DI VERCELLI, DIPARTIMENTO DI URBANISTICA 135
5.2.3 AMI, MOBILITY AGENCY OF FERRARA 136
5.3 IDENTIFY AND ORGANIZE NEEDS 138
5.4 DEFINE FUNCTIONALITIES 141

6 SERVICE DEFINITION FOR THE GEOPKDD PROJECT 151

6.1 VALUE PROPOSITION 152
6.2 DISTRIBUTION CHANNELS 156
6.3 CUSTOMER RELATIONSHIP 157

6.4 VALUE CONFIGURATION 158
6.5 CORE COMPETENCY 160
6.6 PARTNER NETWORK 161

CONCLUSIONS 163

BIBLIOGRAPHY 166

FIGURES AND TABLES 168

Introduction and structure of the work

This thesis is about the work done inside a consultancy project ordered by Wind Telecomunicazioni S.p.A. to Consorzio ELIS in Rome. This project concerns Wind's participation to the European GeoPKDD project, whose acronyms stays for Geographic Privacy-aware Knowledge Discovery and Delivery. The contribute required from this work is to generate and define a new service concept that exploits the technology developed inside the project and that can be a profitable new business area for Wind itself.

The first chapter offers a general overview on the new service development models, with particular attention to the main features that a technology push model, such as the one faced in the applicative part of the work, requires with respect to a more traditional market pull one. This part of the work ends with a deepening into the main strategic decisions that a firm has to take if it wants to undertake the development process of new technology push services.

Then, in the second chapter, the different phases of the creation and the definition of a new service are described, together with the theoretical arguments supporting them. The logical sequence of activities starts with the generation of ideas, that aims to create new concepts for final services to be developed, passes through the screening and the scoring of these ideas and ends with the final selection of the best concept. The final service definition phase is also treated, which includes the client for which the service is thought and the needs that it will address.

Because of the technology push nature of the development process, the purpose of chapter 3 is to shed some light on the technology at disposal in order to better understand the context from which this work started. Therefore, an overview on the GeoPKDD project is proposed, with particular attention on its objectives, its scope and background. The project is also examined from a technological critical point, highlighting the current technological constraints and the research fields on which the project must work. The role of this thesis is contextualized throughout this chapter, showing the various connection points with the rest of the activities which make up the project itself.

After this contextualization, the actual applicative work starts from chapter 4, where the work done during the collaboration with Wind is presented. The first activity is the generation of new ideas, that was made through a brainstorming session together with a market scouting of existing applications. Firstly, through a screening phase, it was possible to identify six highly potential service concepts from the creation phase; then, through a successive scoring process, the service that was finally chosen is a traffic management tool that addresses important needs expressed by mobility agencies all around Italy, along with other potential customers.

Furthermore, chapter 5 deepens the service concept selected in the previous phase and provides the definition of the single functionalities that have to be implemented in order to satisfy market's needs. Firstly a market analysis is provided, in order to identify which are the target clients of the service and which are their main features; after this, direct contacts were created with some of these potential clients in order to understand the actual needs to satisfy. These needs were then turned into technical functionalities through the support of the House of Quality tool, which gave the list of functionalities to implement as an output.

The work ends with the definition of the service in terms of added value to the customer, possible channels used to deliver the service to the final user, value chain configuration, partner network, relationship with the users and core competencies required for the delivery of the service. The model finally proposed identifies WIND as a content provider, which provides to its customers a web-based service that bundles the software package developed by the GeoPKDD project together with WIND's aggregated user data for specific regions of interest.

A conclusion then sums up the results that were achieved by this work, the value added provided to Wind and to the GeoPKDD project, and an insight in the potential developments for the service proposed.

1 New technology push development models

1.1 New Service Development

The ability of a firm to continuously renew its portfolio of services and products is becoming an essential element of competitiveness in the current context of markets in dynamic evolution. Technology comes to life and dies in a very fast way and follows a specific life cycle: each company has to be aware of the limits of the technology adopted and to invest in the renewal of its set of services or products, in order to follow the rapid evolvements of the market.

Therefore, the development of a new product or service can be essential for the survival of a company in the competitive market. The word “new” referred to a product or a service can correspond to several types of situations:

- new inventions: these are services or products completely new both for the company and for the market. They represent a great risk for the firm since they are not based on the company’s know-how;
- new for the firm but not for the market: they represent the entry of the firm in a new segment of market. The risk is mostly on the commercial side;
- new for the market but not for the firm: they represent the enlargement of the portfolio of products or services with reference to a line previously commercialised. The risk is mainly technical;
- improvements of previous services or products: they represent the change of a product or service already commercialised by the firm. The risk is limited because the company uses its distinctive competence for the innovation process;

- new positioning: it refers to products or services previously commercialised that the firm places again in the market because they can now perform functionalities that were unknown before. Also, in this case the company uses its distinctive know-how, so the risk is low.

There are several models for a new service development according to the peculiarities of the firm and of the situation in which the service is brought out. The main situations are [20]:

- market pull: the firm starts the development of the new service because a new market opportunity is identified. Then, the firm identifies the best available technology requested to satisfy users' needs. In this situation, the market itself pulls the development decisions;
- technology push: the development of these services starts from the individuation of a technology's opportunity. Usually, the company owns a technology and wants to create a final application that uses this available technology. The ownership of the technology pushes the development because the company searches for market opportunities that are created by the innovation at disposal. The main difference from a market pull situation is that, instead of starting analysing the market in order to find a new opportunity, the firm has to find only the market opportunities that can use the owned technology;
- platform-based: the new service or product is developed around a technological base that already exists, that is the technological platform. This technological base provided by the platform is used in as many services as possible because it requires a great initial investment. Platform situations are in some sense very similar to technology-push ones because, also in this case, the team begins the development effort with the assumption that the final concept will embody a particular technology. The main difference is that, in this case, the platform already demonstrated its successful

potentialities and so the risk is lower;

- process-intensive: this situation fits better with the development of concrete products than with services and it refers to those cases in which the characteristics of the final product are strongly dependent on the manufacturing process from which it is generated. Therefore, in these situations, the design of the product can not be separated from the design of the process and, usually, these types of products are sold in massive quantities;
- customized: these services are strictly adapted to the clients' needs. In these cases, the user need analysis phase is crucial in order to identify which features have to be implemented in the final service and to completely satisfy the user's requirements.

A synthetic overview on the several development models is presented in the table below [15]:

	Generic (MP)	TP	Platform Products	Process Intensive	Customized
Product Development Description	The firm begins with a market opportunity, then finds appropriate technologies to meet customer needs.	The firm begins with a new technology then finds an appropriate market.	The firm assumes that the new product will be built around an established technological sub-system.	Characteristics of the product are highly constrained by the production process.	New products are slight variations of existing configurations.
Distinctions with respect to the generic MP process		Planning phase involves matching technology and market. Concept development assumes a given technology.	Concept development assumes a technology platform.	Both process and product must be developed together from the very beginning, or an existing production process must be specified from the beginning.	Similarity of projects allows for a highly structured development process.
Examples	Most sporting goods, furniture, tools.	Gore-Tex rainwear, Tyvek envelopes.	Consumer electronics, computers, printers.	Snack foods, cereal, chemicals, semiconductors.	Switches, motors, batteries, containers.

Figure 1-1 - Types of Development Models

Along with these different types of development, in the table it is also explained how the several development processes differ from the generic market pull model.

Product and service development models have evolved through the years since their conception. Up until the 1960s, the dominant model of innovation was a linear technology push type as shown below [19]:

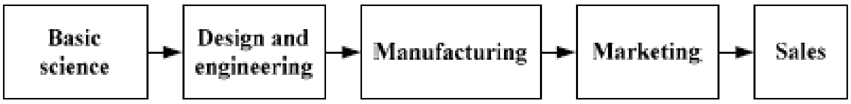


Figure 1-2 – Linear technology push model

. This linear technology push model started with some basic science or technology, went through a design and engineering phase, and was then produced and sold.

As competition increased during the 1960s, a greater attention was paid on the role of the marketplace in development and a linear market pull model became the dominant model of innovation [19]:

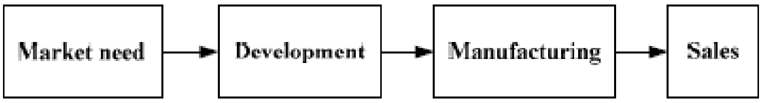


Figure 1-3 - Linear market pull model

This linear market pull model started with a market need, then development occurred using existing technologies to meet the market need, and at last the product was made and sold.

At the time when the linear market pull model became more dominant than the technology push model, significant literature then began to focus on a debate concerning whether market pull or technology push was better for initiating true innovations leading to commercially successful products. During the 1990s, some authors started stating that the best development method should be a combination of market pull and technology push methods. The predominant model of these years was a sequential combination of the market pull and technology push models as shown below [19]:

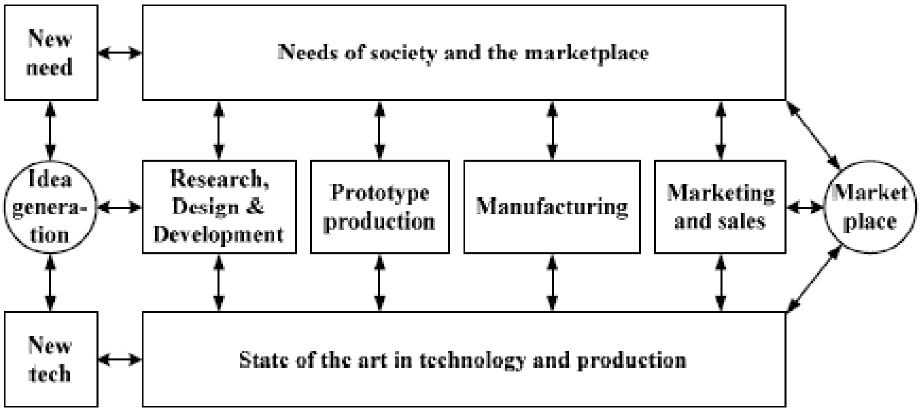


Figure 1-4 - Combination of technology push and market pull model

. This sequential model used both new needs that were recognized in the market and new technology as sources for product development of successful products.

1.2 New Technology-Push Services

Firms that do better than their market competitors often derive their success from innovation and, in many cases, this innovation is technology based. As a matter of fact, the creation of new products, processes or services is recognized to be the main source of competitive advantage, and this innovation can be enabled only by technology.

It is common knowledge that the competitive context is rapidly changing and that technology is the main engine of this change. The number of technological changes is very high and firms have the need to own the ability to exchange technology in a very quick way. In this kind of context, technology and technology innovation are key success factor for the competitiveness of a company in the market today.

Technology can be defined as a body knowledge, tools and techniques, derived from both science and practical experience that is used in the development, design, production and application of products, processes, systems, and services [3].

Through this definition, it is possible to underline some key concepts related to technology:

- it is embodied not only in the final application, but also in the processes or methods needed to generate the new innovative service or product;
- it is the mean through which scientific and engineering knowledge is used into a practical and usable results;
- it deals both with science and technique, but with some differences: science refers to general knowledge about natural and social phenomena; technology refers to knowledge used to solve specific problems; technique, at last, refers to specific knowledge used to solve problems case by case. Science, technology and technique are therefore three different kinds of knowledge that differ only based on the level of generalization. Firms usually deal with technologies and techniques more than with science since they are interested in practical results and solutions to specific problems;
- there are two different ways of developing technology: the first one starts from scientific knowledge and brings to practical applications; the second one, instead, starts from empirical knowledge and brings to its theoretical generalization;
- from a business point of view, technology can not be interesting per se but only when it is closely linked with innovative objectives;
- within a firm, there is a set of tacit knowledge that can be a key factor in achieving results: for example, knowledge embodied in people’s mind, firm’s culture and values and so on.

Innovation can be considered as composed of two parts [5]:

- the generation of an idea, or invention;
- the conversion of an invention into a business or into an useful application.

In other words, innovation can be defined as the sum of invention and exploitation:

- innovation is related to the creation of new ideas;
- exploitation is related to commercial development and focuses on specific objectives, actual results, and vast diffusion of the technology-based outcomes.

Innovation can be divided into two different types [2]:

- minor or incremental innovation, that consists in a modification of an existing entity;
- major or radical innovation, that consists in an entirely new entity.

In the dynamic and changing context in which companies have to compete today, technology innovation assumes the following key characteristics:

- cumulativeness: what a firm will be able to do in the future strongly depends on what is has done in the past. Past experience of a company creates a specific know-how, usually embodied in tacit knowledge of people and procedures, that plays a central role in the innovation process;
- specialization: firms need to concentrate their efforts on few core disciplines in order to keep their competitiveness;
- geographical division of technological research: a division of work in the creation of technological knowledge is needed in order to favor cumulativeness and specialization of innovation activities in limited areas;
- uncertainty: since the technological progress accelerates, the uncertainty of innovative process is continuously growing;
- technology integration: technology innovation is not only the result of in depth studies in one specific field, but it can be often achieved

integrating different knowledge from different fields.

Technology push projects have some features in common and these characteristics associated with technology push models can be grouped into the following categories:

- difficulty: as in market pull, the risks in technology push processes come from several different sources including technology, market, financial, and more, but those risks are more intertwined than in market pull. In stable markets, it may be possible to separate the technological and market risks but that is not so in an evolving market that accompanies the introduction of a radical technology. The complexity of technology push is additionally increased because currently everything is set up to look for exactly what clients are looking for rather than to assist in the process of finding crossover, innovative applications. This causes the matching of customer needs to the technology capabilities difficult which translated to a difficulty in identifying product or service specifications needed for development. Technology push has a higher risk of failure than market pull with 90% to 99% of all technology push development processes failing [2] and even after the technology is relatively proven, the risks are significant;
- company size and resources: while large companies have the needed resources to invest in long term technology push initiatives, they usually lack the incentive to do so. It is often difficult to get support for radical projects in large firms where internal cultures often push efforts toward more low risk and immediate gains. Furthermore, related to the size of a company is the aptitude for communication within a company: medium and small companies have the special advantage of being smaller and more dynamic, with the highest incentive to develop innovative projects. However, because technology push development cycles are so long, expensive and risky, small companies generally lack the resources to be successful;

- benefits of technology push: though 90% of new projects from 1989 to 1993 were market pull types of products, they only constituted 76% of the new profits, while technology push projects accounted for 24% of the profits [2]. Though technology push may have a higher risk, they also have a higher profit picture which may justify the project despite the additional risk. Profit can be gained through technology push methods either by developing new products or by identifying potential applications outside of the company's scope and then licensing that technology. Customers can only ask for services that are better than what they already have, because they don't know the possibilities and won't be able to imagine the revolutionary ones.

Many authors identified characteristics or factors that have been correlated with successful technology push development. The most mentioned factors among the several authors are the following ones [14]:

- there is focus on the final user;
- internal and external networks were used;
- there is management support;
- the development team is dynamic, motivated and talented;
- a combination of market pull and technology push is used;
- the market is developed during or directly after product or service development;
- the technology offers a clear advantage;
- alternatives were carefully examined.

1.3 Technology push development models

Let's now see how a technology push development process can be modeled. A first example of technology push model is presented in the figure below [22]:

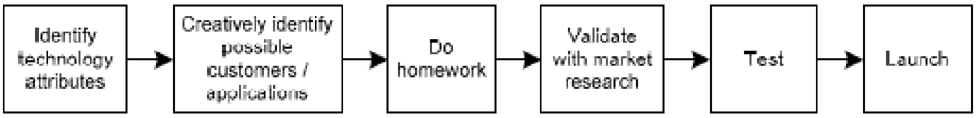


Figure 1-5 - Technology push six phases model

This model suggests to start first with the identification of technology attributes before analysing which are market's needs. A similar model is proposed below [23]:

STAGE in Technology Movement	RELATED ACTIVITY in New Product Development Process
Disclosing Technology	
Linking Technology With Needs	1. Initial Screening 2. Preliminary Market Assessment
Assessing Technology	3. Preliminary Technical Assessment
Matching Technology with Functional Needs	4. Detailed Market Study
Refining Technology for Specific Needs	5. Business/ Financial Analysis 6. Product Development 7. In-house Testing
Preparing to Launch into the Users' World	8. Customer Tests 9. Test Market 10. Trial Production 11. Precommercialization Business Analysis
Managing a Technology Over Its Life Cycle, especially at Introduction	12. Production Start-up 13. Market Launch

Figure 1-6 – Stages of technology push model related to market pull model

In this case, the aforementioned model is compared to a market pull model and it is concluded that there is little difference between the two processes. As a matter of fact, in this comparison there are far more similarities than differences between the corresponding activities. The only relevant difference is that technology push process starts with the disclosure of technology in order to understand its opportunities before gathering needs from the market. It is important to underline that, though the author describes the technology push and market pull models as consisting of stages, he is careful to point out that any development process cannot be considered strictly sequential.

Furthermore, an eight step descriptive model to technology push has been identified [24]:

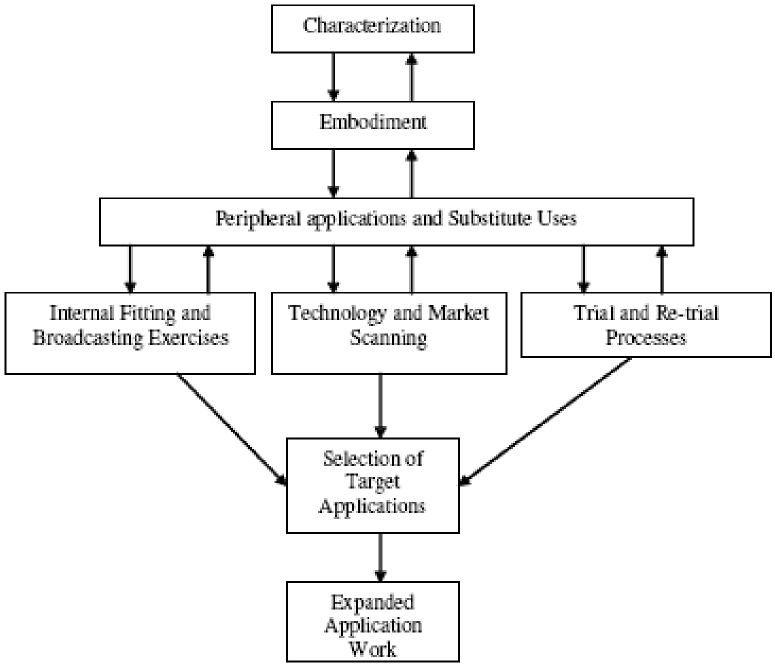


Figure 1-7 - 8 steps model

This is the first technology push model significantly different from the market pull model. It includes first a characterization of the technology and then an embodiment into an application; after this embodiment, other possible applications or uses are explored through several techniques such as market scanning and broadcasting exercises. Then, the final application is selected and the first seven steps are continually iterated. In this model it is not explained when the service is ready for launch, leaving some ambiguity as to when the development process is complete.

Some authors define a process to adapt the traditional market pull model to a technology push model. This step, alternatively known as technology transfer, proposed a way to transfer technologies to a push development by the use of a

Technology Application Selection (TAS) process, that is composed by three phases: Technology Characterization, Application Identification, and Application Evaluation [25]:

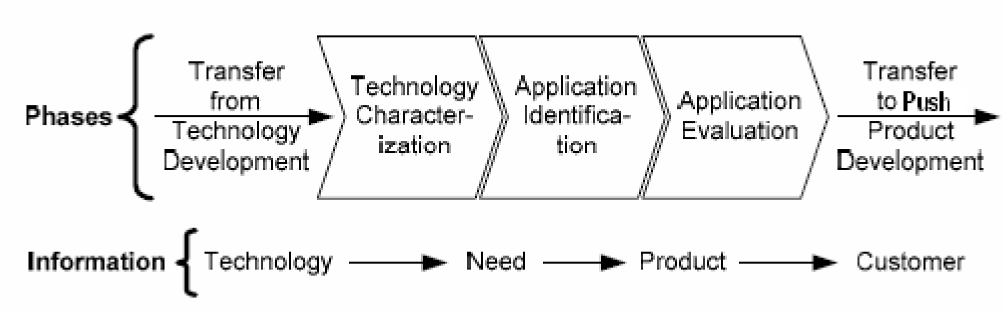


Figure 1-8 – TAS process

In brief, this model means that, once the technology is sufficiently mature, it is then characterized and applications are identified and evaluated. Subsequently, a generic market pull approach is prescribed which requires the technology to be included.

Up to this point, many similarities between market pull and technology push models expressed by many authors have been presented. However, other authors have proposed that market pull models are deficient when used for technology push development. These deficiencies can be grouped into four areas:

- the method of gathering market information: when market pull methods are used to gather market information for a technology push project, it produces unhelpful or even damaging results. Developing an entirely new category of service is risky and, in this situation, market pull model can be used only to define market requirements and to create specifications;
- the use of informal networks: these networks are individuals spanning across industries, companies and functions within companies. These informal networks can be useful to gather and disseminate information in a technology push process, while this critical issue is not even considered in market pull models;

- the use and evaluation methods of prototypes: how developers interact with future customers in the use, evaluation and feedback methods in technology push development is significantly different from market pull situation. As a matter of facts, prototypes in technology push development help in preparing the market for the technology while also generating new concepts. In the market pull model, instead, only after prototypes have been evaluated by customers can real customer needs be identified with confidence and there are no methods to develop prototypes without customer requirements;
- the goals and evaluation of projects: while it is the optimistic goal of every designer to hit the customer needs right on at the first try, in the event where the customer isn't even identified yet, it is realistic not to expect success on the first, second or even the third try.

1.4 Strategy in technology push models

The actual acceptance of the connection between technology and strategy was achieved in the early 1980s, when the first works dealing with technology as a strategic variable started to be published [1]. Before then, technology was considered as a separate subject from strategy and it was treated as a mere implementation issue: this means that a firm used to determine its strategy and technology had the only task to put strategy into reality.

In a dynamic context, decisions about technology strategy are vital for the survival of the firm in the market. The framework proposed defines three categories of decisions: selection, timing and acquisition mode [2]:

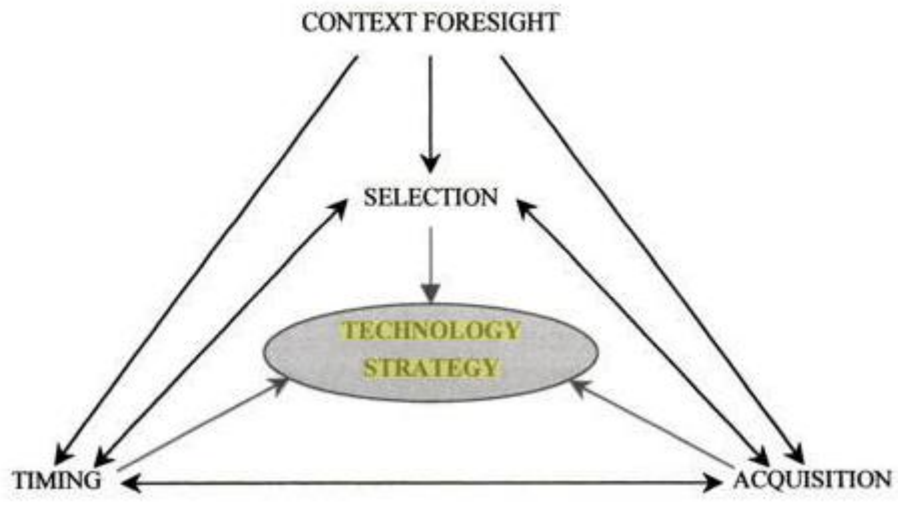


Figure 1-9 - The dimensions of technology strategy

Context foresight

It is the process by which companies identify the key characteristics of future competition. It is the result of two different analysis:

- externally driven analysis: it is made firstly through the identification of how the market will be shaped in the future and how customers' needs will change, then through the identification of the best technologies able to satisfy market's needs;
- internally driven analysis: it is based on the identification of the technological skills available within the firm and of the application that can exploit these potential competences.

Selection

This selection of the technology should be based on these main principles:

- the coherence with the business strategy, which can be differentiation, cost leadership, cost focus and differentiation focus. Therefore, technology strategy should support the kind of competitive advantage that the company is trying to achieve;
- whether the technological change is desirable for the company: the

changes in the industry field have to be accurately considered because sometimes, although a technological change can create advantage for the firm, this advantage can reduce the profitability of the whole industry and therefore can reduce the success of the firm itself in long term.

- relevance: a firm has to evaluate how much the technology is able to create value-added for the final customer and can be used in a final application to satisfy clients' needs;
- risk, that can be of three types: technical, that is the risk that the technology does not achieve the desired level of performance, or that the development program does not end by the time requested; commercial, that is the risk that the associated application fails in the market or has a low economic return; financial, that is related to the amount that has to be spent to develop the technology and to bring it to the market;
- appropriability: it reflects the extent to which the technical knowledge can be protected against imitators with the raise of entry barriers;
- interdependencies: innovation is not only a breakthrough in a specific field but can be the fusion of technologies coming from different areas, and technology value can be related to the number of applications to which it is associated;
- option creation: even if sometimes technologies may not lead to marketable applications, the experience accumulated investing on them can allow new technological directions and can increase the ability for further technological developments.

Timing

The timing of introduction of a new technology involves the strategic decision of being leader of the innovation process or follower. A firm has to consider the sustainability of being leader, that depends on several factors:

- the source of the technological change, that can be within the industry or external. If this source is internal, the technological leadership can be easily maintained while, if it is external, other firms can have access to this source and the leadership is not easily sustainable;
- the advantages related to development. Sometimes a company can have advantages from the development of a technology such as scale economies or higher R&D efficiency, and this makes the leadership more sustainable;
- technological competencies. If they are unique with respect to the competitors, the leadership is easier to maintain;
- the rate of diffusion of technology. Imitators can increase their competences about a technology learning by leaders through press conferences, scientific publications, consultants and so on. However sometimes this is not possible because of the existence of patents and this makes the technology leadership more sustainable.

The main advantages of being the first movers are related to:

- reputation: this can be done only with the help of marketing resources and its advertising activities;
- positioning: first movers can decide where to position themselves within the business, forcing their competitors to other less favorable positions;
- switching costs: first movers can create an high entry barrier in the new technology creating high switching costs for the clients;
- distribution channels: first movers can select the best distribution channels and create a an exclusive access to them;
- learning curve: first movers starts in advance the learning curve with respect to the competitors;
- access to input resources: first movers can take advantage from the access of resources, especially when they are scarce;

- standard definition: first movers can set the standard of the whole industry;
- institutional barriers: first movers can protect themselves from competitors through the raise of institutional barriers such as patents, privileged relationships with government and so on;
- initial profits: at the beginning of the life cycle of a new product or service usually prices are higher and this means that high profit can be easily made.

On the other hand, first movers have to bear these disadvantages:

- pioneering costs: first movers have to sustain several costs such as customer training, infrastructure building, government approval and so on;
- uncertainty: at the beginning of a new project's life, market demand is always uncertain;
- changes in customers' needs: when a new project is developed, users' needs are not so clearly define and can easily change;
- imitation: first movers have to face the possibility of imitation by followers at low cost.

Acquisition

It involves the defining whether to develop the technology internally with owned resources, or to buy it from outside. The main variables affecting this strategic choice are the availability of external sources, time required to develop the technology, costs, technological risks, commercial risks and so on.

It is important to underline that sometimes firms can generate and license new technologies even if they are unrelated with the firm's objectives only because they can create a commercial value and can become a real business. This decision should be taken when license allows to:

- exploit the technology which, otherwise, would remain not exploited;
- access markets that, otherwise, would be not available;

- introduce a new standard more rapidly;
- raise entry barriers;
- have higher profits.

The types of technology strategies can be defined into a technology and application matrix, as shown below:

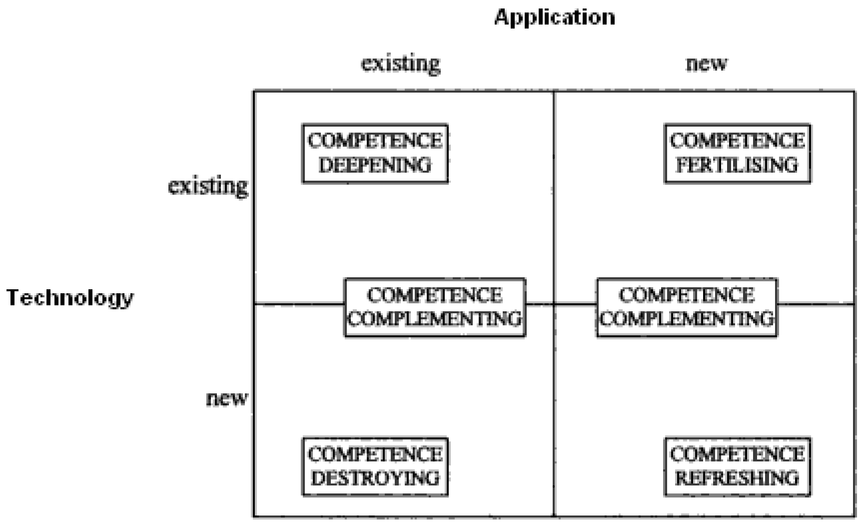


Figure 1-10 - Types of technology strategies

- competence deepening: it requires the investment on the technologies that are critical to the current company's strategy. This strategy involves the concentrating of investments on the deepening of the current base of knowledge and the reinforcing of a firm's technology base. Undertaking such strategy can bring to maintain the knowledge gap with respect to the competitors;
- competence fertilizing: it means to invest on technologies already available within the company which show strong potential to create new applications. This development can also bring to the cross-fertilization of other existing applications if they use the same technological base. From a business point of view, this strategy can lead to problems in getting familiarity with the market of the new

application and, to solve this problem, alliance or joint venture can be useful; from a technical perspective, instead, it can mean to reinforce the current competences through providing a wide range of new opportunities associated with it. Since these competences are already internal to the firm, the developing of these new application are likely to be an internal R&D task;

- competence complementing: it means to acquire new technologies to be integrated with the existing ones in order to open new market opportunities. It relies upon the integration of available technologies with new ones in order to generate new applications or new modes of doing existing applications. The technological risk is reduced since the firm is dealing with the combination of the new technologies with others already known; the commercial risk is also low since the company is already familiar to the application market. It can be an effective way to gradually change the firm's technological knowledge;
- competence destroying: a strategic decision can also point out that certain future technologies may erode the set of knowledge required for existing applications, so that these new technologies are likely to prevail in long term against the existing ones. The earlier the company recognizes that a competence destroying a set of existing technologies is emerging, the earlier it can renew its competences. The company, in this case, has to create a new technology base for new applications in order to keep surviving and it can be made through alliances or joint ventures, but this could reduce the level of appropriability;
- competence refreshing: this investment is aimed to the acquisition of new technologies which have the potential to generate a set of new future applications. Undertaking this kind of strategy may be highly risky as they involve new technology base and concern new applications. If the new technology base is promising, the acquisition mode is the one of acquiring companies that have

know-how in that field. In the case the technology base shows a strong potential but it is still in an embryonic phase, internal ventures or venture capital investments could be appropriate in an initial phase, while an internal R&D can be undertaken in later phases.

The central element of this matrix is the defining of a strategy trajectory, which means that a continuity in technology actions should be found. The trajectory refers to a long term program according to which a company should start investing in competence deepening and fertilizing, in order to acquire familiarity with the new application and the new market. Then, competence complementing process helps in acquiring technologies that can be critical in the future, and this makes easier to realize a continuous learning process through the refreshment of the current competence in a further phase. This refreshed competence becomes the new current base and the cycle restarts.

2 New service conception and definition

In the applicative part of this work, the firsts steps of a technology push new service development will be faced and explained, with reference to the case of the European GeoPKDD project. What will be examined is the process that starts from the analysis of the technology opportunities and ends with the definition of the functionalities to implement in the final application. In particular, the logical stream of the activities within the whole development process that concerns this work is the following one:

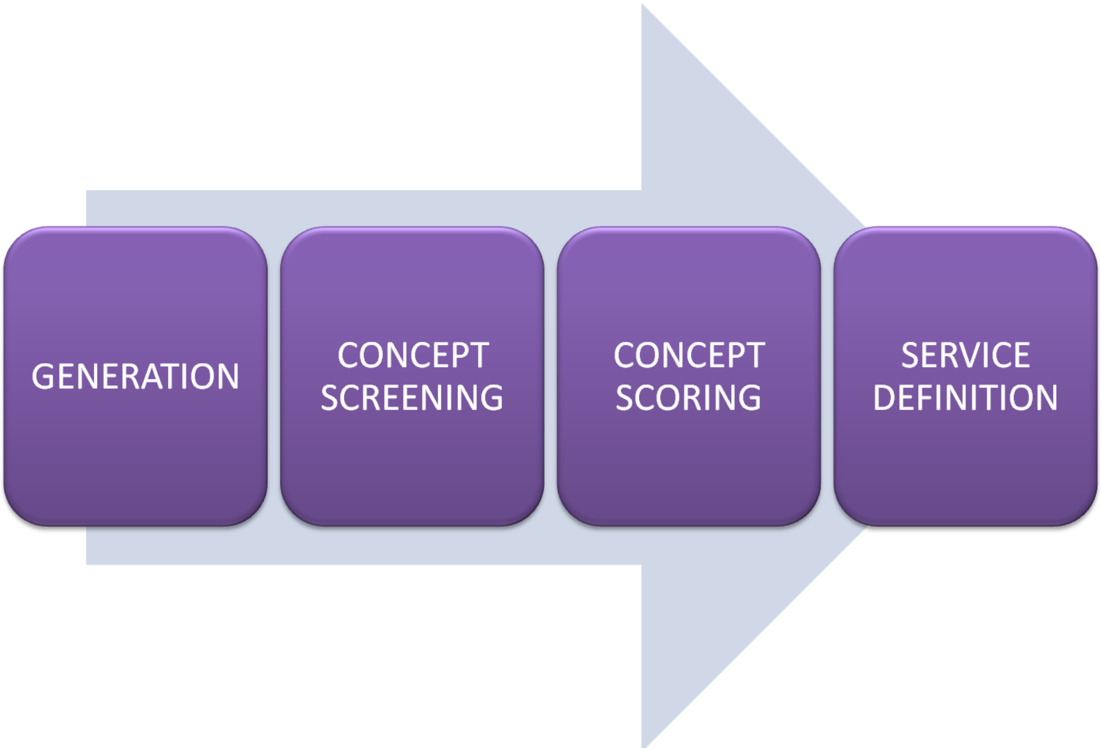


Figure 2-1 - Logic sequence of activities

The main issue is that the new service that this work aims to bring out is technology based, that implies that the process starts with an in depth analysis of the technology opportunities and constrains before studying the market and seeing how it can take advantage of the innovative researches.

In this section, the logical steps of this methodology will be explained in order to see how to perform the process that starts with the generation of ideas and ends with the definition of the new technology based service; a general theoretical basis about these phases will be given, together with the explanation of the main techniques that will be used in the applicative case in order to well perform them.

2.1 Generation

One of the first and most important phases of each development process is the conception of the innovative idea that will be implemented in the new service or product. Therefore, ideas generation activity is a crucial point of the development process and it should end with a large number of concepts of services, out of which the final concept will be selected to be used in the further steps of the process.

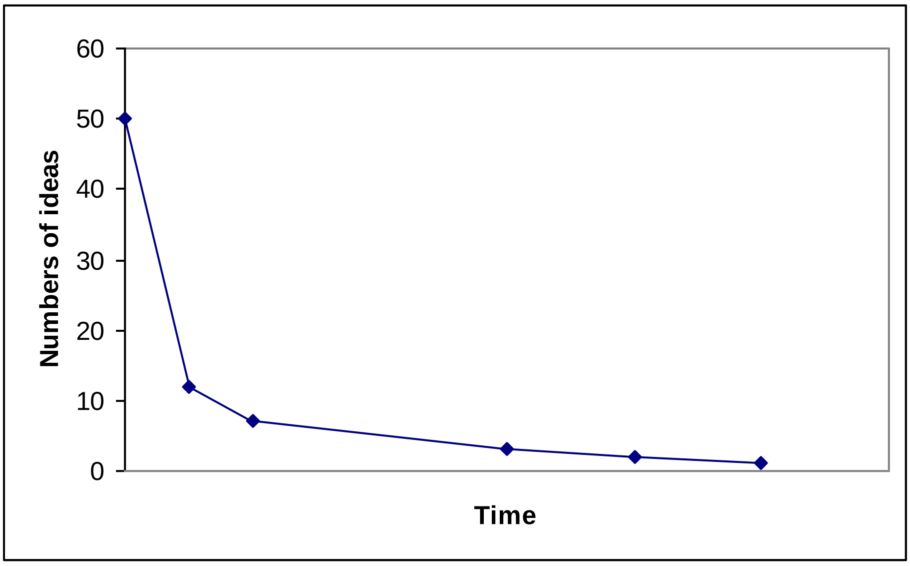


Figure 2-2 - Number of ideas

Ideas generation has to be performed in order to guarantee that the entire set of possible solutions has been explored in a complete way and that all the possible aspects of the problem have been taken into account.

Specifically, there are some common threats that can be found in the concepts generation phase, such as:

- only one or two alternatives are taken into account, usually the ones brought out by the more influent members of the development team;
- the utility of concepts used within other companies for similar products or services is not taken into account,
- entire categories of solutions are not considered;
- several promising solutions are not taken into consideration because their integration is not easy to achieve.

The generation activity is not expensive and takes a short time with reference to the other development phases; however, it is important to perform it with attention in order not to compromise the results of the entire development process.

As a matter of fact, this phase absorbs a small percentage of the development costs but, at the same time, it has effects on all the other activities of the development process. As it is shown in the figure below, the possibility of influencing the costs has its maximum at the beginning of the development process, while information about the costs is very scarce:

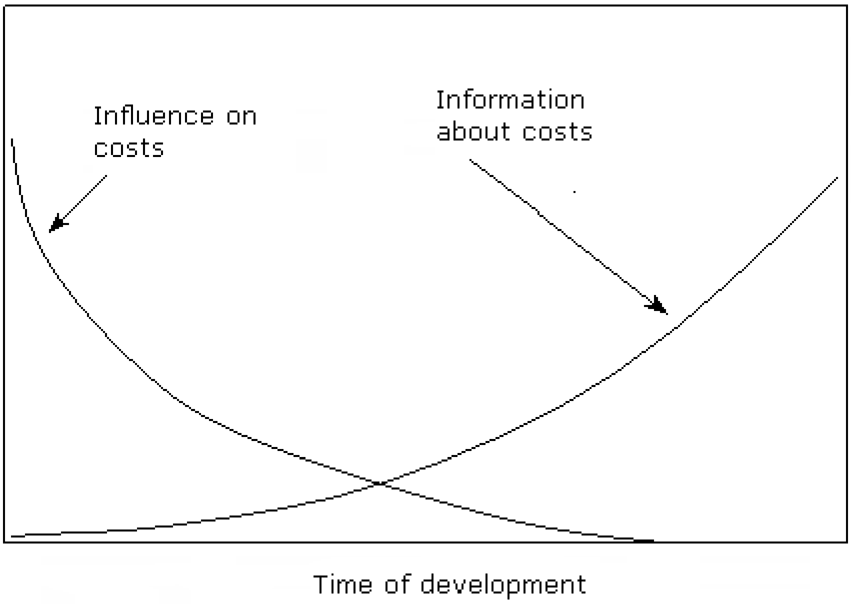


Figure 2-3 - Costs during development

Ideas generation is a crucial phase in the development of a new product or service. The sources of new ideas can be of different nature, both internal and external to the firm. Some of the sources for the creation of ideas are:

- clients: ideas can come from direct or indirect requirements of the user through complaints, after-sale activity and so on;
- competitors: many ideas can come from the analysis of competitors' developments, tests performed, investments made, patents, and so on;
- supply chain: suppliers, distributors, retailers and others actors involved in the supply chain can offer interesting ideas especially because, sometimes, they have a more direct contact with the final user;
- lead users: they are the advanced users of the service or product and so they can feel the market's needs in advance with respect to the rest of the users. They are a great source of innovating ideas, especially in the technical field;
- experts: these can be consultants, university professors, employees of competitors' firms and so on. Their contribution can

be determinant especially with respect to technical innovations;

- technical journals or publications;
- internal sources: many ideas can come from R&D researches or from internal intuitions.

There are several methods that help to generate new ideas, both for collective and individual reasoning task. Let's see an overview of these two ways of innovative thinking.

In collective situation, the focal hint is to keep separated the generation phase from the evaluation one: if it doesn't happen, each trial to propose a new idea can be sunk by critics that prevent any deepening. This behaviour has the effect of raising the inhibition of people and to keep thoughts within rigid schemes.

The collective generation of ideas can product a radicalization of each one's positions, that implies that a participant of the generation session defends his own idea in order to impose his own role inside the group. This can create the bad effect that an idea is selected only because of the good dialectical ability of the one who proposes it. Therefore, it is important to separate ideas from people and to remember that the goal is to create innovative ideas, not to demonstrate the own personal ability.

The most common techniques for collective generation of ideas are:

- analogy with similar cases: the development team can find new ideas searching for solutions similar to the existing ones that can be used for a similar problem;
- reverse engineering: the team can analyse a competitors' product or service in order to imitate it and, possibly, to understand how it can be improved;
- brainstorming: it is aimed to bring out from a group of people a set of creative and innovative ideas for a specific context. People can say ideas in a free way and each idea has to be taken into account, without any judgement or censorship;
- 6-3-5 method: it is also called "written brainstorming method". According to this method, six persons have to meet and each of

them has to propose three solutions. Then these ideas are put in the middle of the table in an anonymous way and there are five further phases in which these ideas are elaborated, modified and improved by the other participants;

- TRIZ, Theory of Inventive Problem Solving: the basic scope of this method is the resolution of the internal contradictions of the problem. According to this method, a matrix has to be drawn in which each cell represents the trade off between two different characteristics of the final product or service. In each cell, up to four principles taken from forty basic physical principles are suggested in order to solve the trade off. This method fits better the generation of a concrete technological product;
- the Six Thinking Hats technique: it is based on the idea that seeing things from different points of view is often a good way of generating new concepts. The technique is designed to help explore a variety of perspectives on a specific subject. In wearing a particular thinking hat, people play roles and put themselves into a particular perspective. The hats are:
    - white: it is neutral and focuses on information available, objective facts, what is needed, how it can be obtained;
    - red: it represents emotions, feelings, intuition; it presents views without explanation or justification;
    - black: it represents the judge wearing a black robe, and it is judgmental, critical, considers why something is wrong and has a logical negative view;
    - yellow: it represents optimism, has a logical positive view and looks for benefits and for what's good;
    - green: it represents the creative thinking, considers possibilities and hypotheses, stimulates new ideas;
    - blue: it represents overview, takes the process under control and is the organizer of the process.

Some of the benefits of the Six Thinking Hats method are:

- it allows to say things without risk;
- it creates awareness that there are multiple perspectives;
- it leads to more creative thinking;
- it improves communication.

Tools have also been designed in order to help the process of idea generation in individual contexts. Some of the most important are:

- mind mapping: it is a technique of visually arranging ideas and their interconnections. It can be used to graphically arrange the links of some central concept or issue with other concepts and allows to create and organize ideas, as well as to communicate readily them. It always starts from some problem or issue which is positioned in the centre and typically it contains words, short phrases and pictures, which are connected to the central concept by lines. Unlike linear thinking modes it stimulates imagination and creativity, by connecting left and right brain thinking. The main benefits of this tool are:
    - simplicity: it is easy to use;
    - associativity: it stimulates the creation of the many links that any idea can have;
    - visualization: it is easy to remember;
    - radiality: it allows to work in all directions;
    - overview: it helps to see the relationships between issues;
- random words: this method consists firstly in establishing a series of free associations correlated to the focal subject about which a new idea has to be generated, and then in selecting another random word and creating new associations related to it. Finally, crosses between the two series of associations have to be created, in order to generate new and innovative ideas;
- lateral thinking: this techniques is characterized by the shifting of

> classical thinking patterns that want people to think in a “sequential” and logical way. A new idea that is the result of lateral thinking is not always a helpful one, but when a good idea is discovered in this way it is usually obvious with hindsight. There are some mental tools that can be used in lateral thinking. The first is “provocation”, that consists in declaring the usual perception out of bounds or in providing some provocative alternative to the usual situation under consideration. Provocation inserts an element of discontinuity in the thinking pattern and tries to provide solutions taken from completely different fields. Sometimes common reasoning pattern is so embedded in people’s mind that it is not possible to see other ideas. Another used tool is “challenge”, that is analysing the way things have always been done or seen, or the way they are, in order not to show there is anything wrong with the existing situation but simply to direct your perceptions to exploring outside the current area.

In this section, an in depth analysis of brainstorming method is offered in order to let the reader understand better the further part of this work in which a real brainstorming session will be presented in the applicative case.

Brainstorming is a semi-structured creative group method according to which a group of people is put together in the same room and an open and informal discussion is encouraged. The idea behind brainstorming is that a group of people can achieve a higher level of creativity through synergy than the sum of the participants separately.

The objective of the meeting and the problem to which the group has to find a solution have to be clarified from the beginning of the discussion. Each member can expose all the ideas that came in his mind in a free way, without the risk of being judged by the other members. No idea can be refused in this step of the process, because also the most nonsensical ideas could lead to the generation of more effective ones. Each idea has to be jotted down on paper with few words of explanation.

Some useful rules that should be respected during the discussion are proposed below:

- set the direction: the situation and the problem have to be clearly described and defined. People have to understand the problem to be solved and the objectives. The group must be kept on track;
- involve everyone: everyone has to be encouraged to contribute and dominating participants have to be controlled. Everyone should contribute and develop ideas, including the quietest members of the group;
- don't overlook the obvious: the obvious solution is sometimes the best one;
- suspend judgment. There are no bad ideas, all ideas are good ideas. Any judgment should be avoided and any criticism should be prevented in advance. The negative consideration of an idea should rather be turned in new improved ideas;
- generate a huge amount of ideas: if the number of ideas is high, the exploration of possible solutions will be much more complete. The focus should be paid for quantity, not for quality;
- encourage cross-fertilization. Participants should combine, synergize, and improve upon other participants’ ideas. People should be encouraged to build on the ideas of each others, creating unlikely combinations and taking each one in unexpected directions;
- encourage outside-the-box thinking. People should be creative., because sometimes the wildest ideas lead to the conception of the best concepts. Ideas that can seem unrealizable should be accepted: solutions that at first sight seem unfeasible could be implemented with the integration of other solutions. In addition, also these kinds of ideas can encourage the group to bring out other useful solutions;
- don't fear repetitions: at different moments an idea can be seen with different eyes. During discussion afterwards, the duplicate

ideas may trigger a different response at a different time;

- fix a numerical goal: the bringing out of an excessive number of ideas can be really tiring for the group. Nevertheless, people can be motivated by the definition of a minimum objective of ideas to propose;
- use graphical means: graphical representations of the problem to which a solution has to be found, such as slides or drawings, could help people in reasoning about possible ideas. Each idea should be displayed in a complete form, without using one-word descriptions, in order to avoid misunderstanding;
- apply the 80/20 rule: at the end of the discussion, people have to look through the list of ideas and circle the 20% that will yield 80% of the results.

All the ideas that come out from the process have to be written on a board visible by all the participants. So, the actual output of the process will be a board full of sentences, each of them referring to an idea proposed by a member.

With reference to the conduction of a successful brainstorming session, some useful hints are:

- use an experienced facilitator, internal or external;
- during the reasoning task, put questions such as “what would be my needs if I were..” or “what if..?”: this will stimulate the group to consider new opportunities and to reflect about the constraints of the problem;
- appoint one person to write down all ideas that come out of the brainstorming session;
- use a flip chart for making the notes. This enables later study and evaluation at the end of the session;
- identify a precise topic to be discussed in order to keep the session focused on the problem;
- no more than 8-10 people should participate in one session. If there

are more participants, then it is better to split up the session and report to each other afterwards;

- encourage an enthusiastic, uncritical attitude among members of the group;
- ensure that no train of thought is followed for too long.

The human brain is divided in two different zones: the right side of the brain controls creative, visual and spatial concepts, while the left side controls logical, mathematical judgmental and analytical activities. That's why it is suggested to divide the time between two types of activity: in the brainstorming session, people have to suspend judgment and come up with ideas by using their right side of the brain; then, switching and evaluating, the team can arrive at the best conclusion using the left side.

The input required for this first step of the methodology is an in depth analysis of the technological background of the applicative case, that is the GeoPKDD project, in order to analyse and evaluate the opportunities and the constraints offered by the technology at disposal. This is necessary in order to have a major awareness of the kind of services that can take advantage of the developed technology.

After this technical alignment, it is possible to give way to the process of new ideas generation. The means through which this generation will be performed are:

- brainstorming session: in this meeting, it is advisable to involve experts of the technological field, in order to gather their competent contribute and to generate innovative ideas;
- market analysis and scouting: this further work is meant to complete the generation phase by gathering information about similar applications and projects that could provide useful starting points for other innovative ideas.

The generation process will be performed taking into consideration three different semantic areas which each idea refers to:

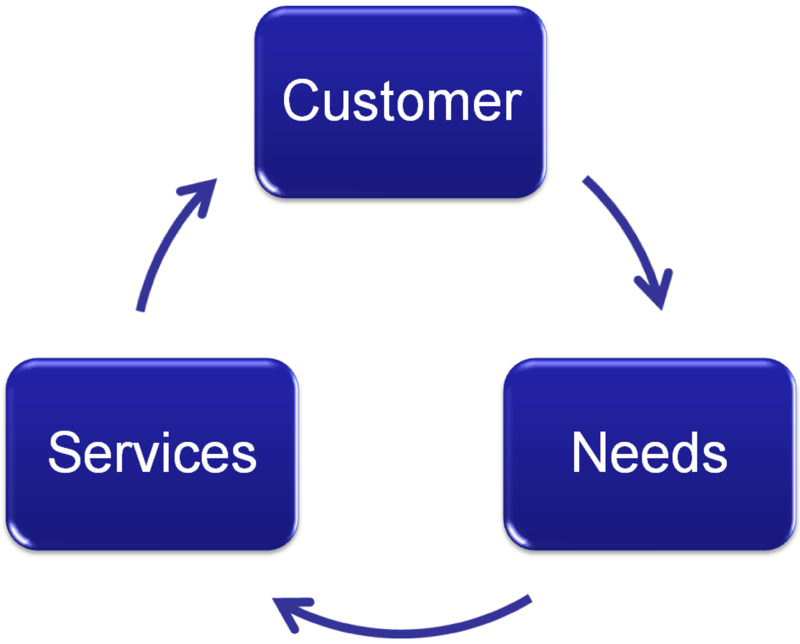


Figure 2-4 - Three semantic areas in the generation process

The figure above is useful to underline that, during the creation of new ideas, it is important to focus not only on services, but also on the clients that the service addresses and on the needs that this customer can have. There could be multiple services that address only one client and, vice versa, there could be one service that addresses more than one client; the same goes for the logic passage from clients to needs. This logical pattern is useful because the identification of possible needs can lead to the generation of new service concepts, that can be the new input for the circle.

The closed loop model shown above is useful also to understand that the generation of the concept doesn't have to start necessarily from the service conception: sometimes it can be more effective to start thinking of possible final users and, from there, considering possible users' needs and final services that can satisfy these needs.

In this way, it will be easier for the members of the brainstorming session to organize their own ideas and to create a good reasoning pattern that ends with the formulation of innovative ideas for final services.

At this point of the process it is really important to have clearly in mind the technology constraints, in order to concentrate the generation efforts only on ideas that can be feasible to implement.

The ideas generated at this level have to be referred to the conception of new services: it means that, at the end of this first step, a list of several new services using the technology at disposal has to be produced.

2.2 Concept screening

This step of the process can have as its main goal to clarify all the ideas that were brought out in the previous phase. The first thing that this method suggests in this step of the process is to undertake a first general elimination of the several ideas, in order not to take into account unfeasible or unrealistic proposals.

In order to easily choose which ideas to eliminate, it can be useful to answer these kinds of questions about each idea:

- is it actually possible to implement this solution?
- would there be anyone interested in such a solution?
- is there a solution in the market that can do the same things better?

This type of simple questions would make the selection activity much easier since they require a Yes/No answer: in fact they divide the ideas between the ones that can go on to the following step and the ones that can not.

At last, all the remaining ideas have to be organized in order to group together the concepts that are similar; so a small number of groups of ideas have to be created, bringing together similar solutions and assigning them a unique identifying name. In this way, the next step of evaluation of the ideas will be faster and easier to realize.

After this, a structured concept screening process has to be undertaken in order to work on the all raw ideas coming from the brainstorming process and to reduce to number of these ideas before the further process of ideas scoring.

First of all, criteria for the screening process have to be selected. Usually each member of the development team provides a list of 10-20 criteria and then the several lists are discussed and put together in order to create a hierarchic order of importance among the criteria. Development costs are usually among these criteria, even though it is really difficult to estimate them in this early stage of the process. In some cases, this kind of criteria is substituted by a more generic evaluation of the complexity of implementation of the concept. At this stage of the process no priority number should be attributed to the several criteria: this usually is made in the further concept scoring when the ideas to be evaluated are fewer.

When criteria have been chosen, a matrix has to be drawn in which each column represents an idea and each row represents a criteria through which concepts are evaluated. It is important that each participant has a complete awareness of the meaning of each concept, in order to evaluate it in a more effective way.

Then a reference concept has to be selected: it will become the concept with respect to which all ideas will be evaluated. It is important to select one of the best concepts in order not to attribute a positive evaluation to all ideas. The reference concept can be, for example, the more obvious idea or an already existing service.

At this point, each concept has to be compared to the reference idea. Each cell has to be filled with a "+" if the idea is better than the reference one with respect to that criteria, a "-" if not and a "0" if the two ideas are indifferent from that criteria's point of view. If there is a concept that is much better than the reference one is avoided to put a "++" in the correspondent cell: a major differentiation among concepts will be made later.

When the matrix is completed, the scores in each column are summed and the final score for each idea is put at the bottom of the matrix. This ranking has to be used to select the concepts that can be the input for the scoring process. For example, the development team can choose to select only the ideas that have a score greater than a certain value. It is important to underline that this ranking has not to be taken in an absolute way: some ideas can have a small general score but a "+" in a crucial criteria, so that it is worth to select them in order to evaluate them in a more structured way in the further step.

The concepts selected can be also improved. For example, a solution can be good but it is penalized by a negative feature: in this case, a little change in the concept can improve it, keeping it different from the others. Another case is when the integration among ideas keeps the positive features ("+") and eliminates the negative ones ("-"). This means that the screening matrix has not to be seen as a static tool, but only as a reasoning support for the decision of which ideas shouldn't be taken into account in the next scoring step.

2.3 Concept scoring

Now that the number of ideas is reduced, each of them has to be evaluated in order to create a final and structured score of them.

Also in this step a matrix has to be built. The criteria can be the same as the ones from the previous stage or also more, since now the evaluation process is more in-depth. It is important to assign a priority number to each criteria, in order to discriminate the importance among them and to make the final score more precise. These numerical weights can be assigned in a subjective way and according to different scales, usually from 1 to 5.

Each idea has to be evaluated according to each criteria. The evaluation now has to be numerical, for example from 1 to 10, and also in this case a reference concept can be chosen in order to make the evaluation simpler.

A weighted sum is then calculated for each column, taking into account the numerical weight of the criteria and the evaluation for each criteria. At the end, a final ranking of the concept is obtained, through which it is possible to select the best concept among the others. It is recommendable to review the votes and to perform a sensibility analysis modifying weights and values. In this way it is possible to estimate how the uncertainty of evaluation can have effect on the result of the process.

There are some useful hints that have to be taken into account during the performing of the scoring process:

- partition of concept quality: the theory is based on the assumption that each criteria can be evaluated independently and the global quality of the idea is the sum of the quality of different criteria. However, the quality of some concepts can hardly be divided in a set of independent criteria and, in the same way, performance with reference to one criteria can be hardly related to the global performance;
- subjective criteria: some criteria are highly subjective and, in this cases, the collective vote is hard to obtain. A solution to this problem could be to ask for en external evaluation, for example the one of a focus group or a group of experts in that field;
- encourage the improvement of concepts: during the discussion about each concept it can be useful to drill down the several contributes and to use these annotation in the phase of combination and improvements of concepts;
- include considerations about costs: costs have to be taken into account in one of the criteria since it affects the final sale price of the service and, therefore, the success of the concept in the market;
- select elements of complex concepts: some complex concepts can be considered combination of simpler ideas that can be evaluated in an independent way.

In the applicative case, three criteria will be chosen with a numerical scale of evaluation that goes from 1 to 10; finally, the best idea will be selected taking into account the combination of values assigned to each criteria.

In order to better visualize the selection process, a particular graphic mean will be used, that is the bubble diagram:

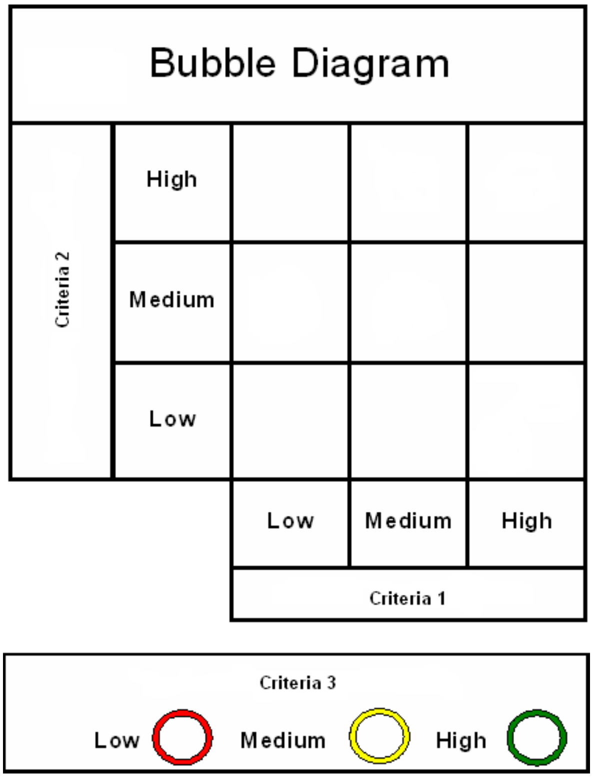


Figure 2-5 - Bubble Diagram

This is a very powerful representation tool since it is able to visualize a three dimensional evaluation scheme, and not only bi-dimensional as usually happens with diagrams.

In order to have this representation, it is necessary to pass from the numerical values assigned to each criteria (from 1 to 10) to a qualitative evaluation of this kind: low, medium and high. This will be made simply creating specific intervals of numbers corresponding to each qualitative judgement, in order to make this passage as objective as possible and in order not to lose information.

The first two criteria are represented on the x and y axis, as shown in Figure 2-5. The third criteria, instead, is visualized with a different colour assigned to the bubble in the diagram: red bubble means low evaluation, yellow medium value and green means high value. In this way, with a single graph it is possible to show the results of the evaluation conducted with three different criteria, in order to simply demonstrate the validity of the choice made. The best idea will be the one located in the top right of the diagram, characterized by a green colour of the bubble.

2.4 Service definition

This is the final step of the whole proposed approach. Once the concept idea is selected, it has to be defined in depth in order to identify which are the functionalities that the final service should implement.

The logic sequence of steps through which it is possible to obtain the service definition is proposed in the figure below:

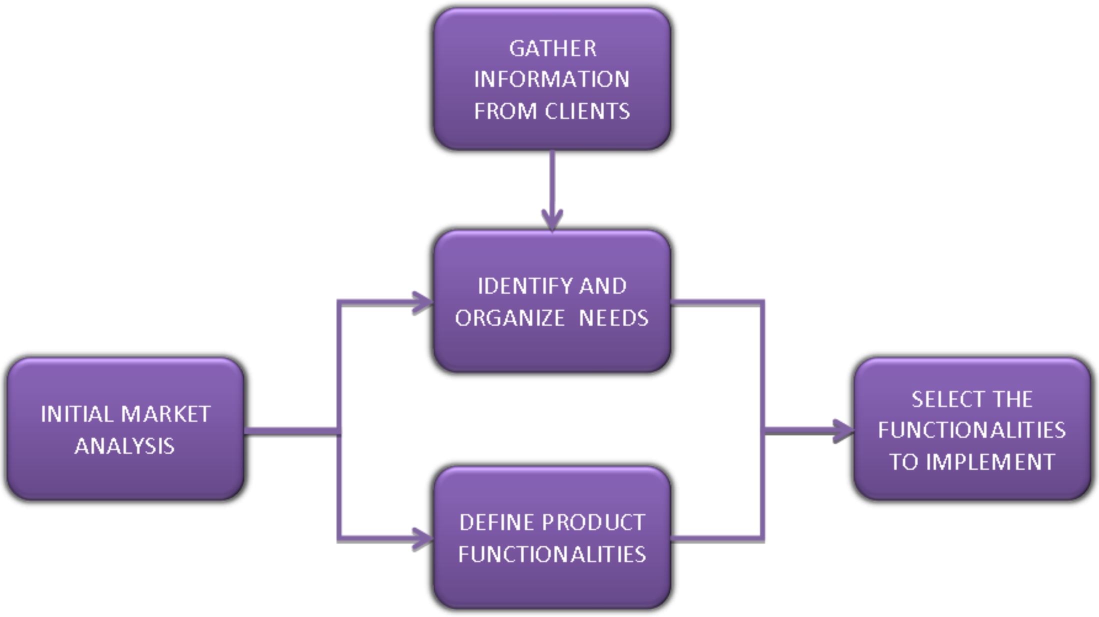


Figure 2-6 - Service definition process

2.4.1 Initial market analysis

An initial market analysis is important in order to understand who are the final clients of the service and which are the main features of the market where the new service will be introduced. As a matter of fact, before directly meeting the user it could be useful to have a look at what the client does and what its research fields are, in order to bring out his latent needs during the actual interviewing. In this way, after the client has freely expressed his needs, the interviewer can actively intervene suggesting other ideas and helping the user to think of new important requirements. The analysis should be concentrated on the whole market and not only on the specific customer, in order to better define important features common to other companies and to gather common needs. This will make the final service more sellable in the whole market and not only to one specific user.

First of all, it is important to understand what is the target market for the idea proposed in the previous step. Different kinds of strategies can be undertaken in a new service development whether the target market is new for the company or it is already existing. In this regard, an important tool used to have a clear picture of the market situation is the Ansoff matrix. Four different strategies correspond to each of the features of the market and of the product, as shown in the figure below:

	Existing product	**New product**
Existing market	**Market penetration** Increase sales to the existing market Penetrate more deeply into the existing market	**Product development** New product developed for existing markets
New market	**Market development** Existing products sold to new markets	**Diversification** New products sold in new markets

Figure 2-7 - Ansoff matrix

With specific reference to the diversification strategy, corresponding to the situation of new product in a new market, four different kinds of sub-strategies can be identified for different cases [1]:

- horizontal diversification: it is the situation of new versions of products or services not technologically correlated to the existing ones that are introduced in segments of market already covered by the firm;

- vertical diversification: the firm improves the quality of its product or service in order to provide a new product that is technologically better the previous one;
- concentric diversification: a growth strategy in which a company seeks to develop by adding new, but related, products to its existing product lines to attract new customers.;
- conglomerated diversification: firm moves in new segments of market with completely new products or services. The company seeks to develop by adding totally unrelated products and markets to its existing business.

After having clearly identified the target market, an in depth analysis has to be conducted in order to identify which are the external opportunities and threats that characterize a specific industry or business. A market research comprises all the factors that are out of the firm's direct control and that have to be taken into account when a company is going to enter a particular competitive field. Factors that can be studied in this kind of analysis are, for example:

- competitors;
- market trends;
- economic conditions;
- existing laws;
- technology trends;
- criticisms;
- changes in markets;
- environmental conditions.

This kind of analysis can put in evidence a set of market opportunities, that the firm should exploit, and possible threats, against which the company should try to defend.

In order to have a more complete scheme, this external analysis can be merged with an internal deepening about company's own strengths and weaknesses. This kind of analysis can take into account several factors, such as:

- availability of internal financial and intellectual resources;
- cost advantages from proprietary know-how;
- inner creativity and innovative capability;
- valuable intangible assets, that is intellectual capital;
- competitive capabilities;
- effective recruitment of talented individuals;
- availability of exclusive information or competences.

Internal and external analyses can be structured together into an unique S.W.O.T analysis, that can be considered a really useful tool to evaluate the appealing of a particular business area for a company.

The acronyms stands for Strengths, Weaknesses, Opportunities and Threats and this analysis can also provide useful hints about how to face the market depending on the internal and external factors previously analyzed:

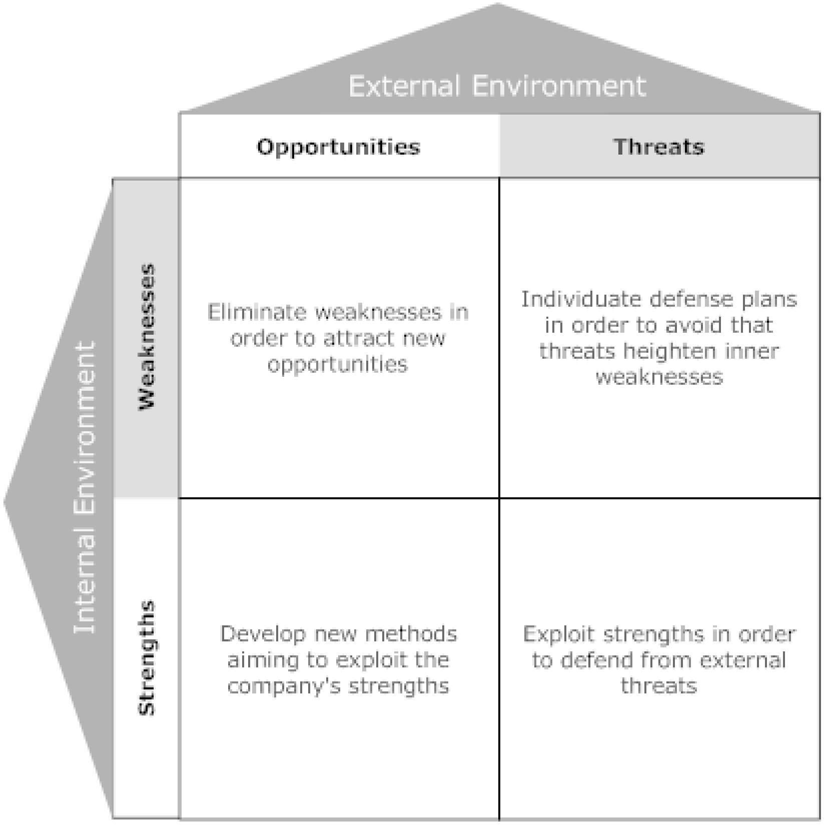


Figure 2-8 - S.W.O.T. analysis

The market research has also the scope to estimate which could be the potential demand for the final service that is being developed. Since the innovative feature of this service, no data about the past demand is available and so only a qualitative forecast method is applicable. Three main kinds of qualitative methods can be used [1]:

- market surveys: they are carried out through direct interviews to possible costumers in order to evaluate their acquisition propensity. Interview is based on a questionnaire that can be structured, if it is composed by a list of formalized questions, or not-structured, if it is carried out in a more flexible way. In the latter case, the interviewer usually uses a check list as a reminder of the focal points to ask;
- opinion of experts: it consists in gathering a group of people that can

be considered experts in the specific field of the service that is going to be developed. A forecast of future demand can be obtained calculating the average of the several indications of the experts or asking the participants to agree in a unique forecast. Through this method previsions are made quickly, at low costs, and different points of view are highlighted. Nevertheless, some disadvantages can be found, that are for example the fact that opinions are less reliable than concrete facts and that bad estimates are taken into account as well as good ones;

- Delphi method: it is an interactive forecasting method involving a panel of independent experts from different knowledge backgrounds. Experts answer questionnaires in two or more rounds and, after each round, an anonymous summary of the several previsions is provided in order to encourage participants to converge towards a unique answer in the further steps. The process continues until a pre-defined stop criteria and then the mathematical average of the final previsions determines the result. This forecasting method aims to obtain the informed judgment of experts, that is something that lies between sure knowledge and mere speculation. It is important to keep the anonymity of the answers in order to prevent influent participants from leading others’ opinions and to pay attention to the selection of the group. This method has the advantage to allow a mutual exchange of competent opinions, but it has some disadvantages: first of all, it is difficult sometimes to discriminate experts from individuals who are only informed; secondly, the convergence of the judgment doesn’t assure the truth; finally, several opinions don’t always have the same value.

In this work’s case, a precise forecast of demand is really difficult to obtain because of the innovative features of the project, but an estimation is required in order to justify the further economical efforts that will be required for the application development. A methodology to forecast the demand involving market surveys is proposed, along with its advantages and constraints.

Three main types of information need to be gathered in order to have this evaluation:

- N: number of potential clients. This measure refers to the total number of users to which the service could potentially address;
- A: potential acquisitions. It's the number of clients that could potentially buy the service if they were informed of its availability. This measure refers to the capability of a firm to invest in advertisement and in informative campaigns;
- P: probability of acquisition for an informed user.

The last measure is the most difficult to obtain since it refers to qualitative and probabilistic estimations. A method to estimate this value is to involve a sample of clients in order to evaluate their interest in the service. First of all, it is necessary to explain the service concept in order to give a clear idea of what it is going to be developed; then, it is possible to ask the client his acquisition propensity that can be of these several types:

- surely he would buy it;
- probably he would buy it;
- he could buy or not;
- probably he would not buy it;
- surely he would not buy it.

It is important to underline that this kind of estimation presents a high level of uncertainty, especially when dealing with a completely new service. Nevertheless, this forecast is supposed to have a good correlation with the real demand.

Having at disposal the answers from the clients, it is possible to estimate P with this empiric formula:

$$P = C_{surely} \times F_{surely} + C_{probably} \times F_{probably}$$

where:

- F_{surely} is the ratio of people that stated that surely would buy the service;

- $F_{probably}$ is the ratio of people that showed a probable acquisition propensity;
- C_{surely} and $C_{probably}$ are calibration constants that are used because people usually overestimate the probability of acquisition; they are calculated on empirical basis and usually assume these values [20]:

  $C_{surely} = 0,4$

  $C_{probably} = 0,2$

At last, after having obtained all these measures, the expected demand Q for the service can be estimated by this simple formula:

$$Q = N \times A \times P$$

Results obtained with this kind of methodology have to be accurately evaluated because forecasts are characterized by a high number of errors. There are several factors that can contribute to the difference between acquisition propensity of users and real acquisitions:

- importance of word-of-mouth: if benefits of a service are not immediately evident, the positive experience made by users can generate new demand for the service. This feature is not usually considered in the estimation of the P measure;
- accuracy in the service description: if the service is not appropriately described to the user, estimations can not be accurate;
- promotion: investments in advertisement and in other forms of communication can increase the demand for a service with respect to the estimated one.

2.4.2 Gather information from the client

After selecting the service, it becomes important to focus on the client that will take advantage of the new service. It is now crucial to understand who will use the service and, in particular, which are his real and actual needs because, at this point of the process, only a generic idea of the service is at disposal: in order to define the concept in detail, it is necessary to get in direct contact with the client in order to understand which are the actual needs that have to be satisfied.

In order to gather information from clients, several methods can be used. The most used ones are:

- interviews: one or more of the members of the group discuss the needs with one single representative of the client. They are usually conducted at the client's site;
- discussion group: a moderator stimulates a discussion with a group of representatives of the client and of the developer in order to clearly bring out what the features of the product must be;
- product use observation: it can bring out important details about user needs, but in this case it is unfeasible because in this project there is not a concrete product yet.

Among the several techniques of gathering information from the client, the one that will be used in the applicative case is the interview.

The questionnaire on which the interview is based can contain different types of questions:

- open questions: they establish only the subject and then the client has the freedom to answer as he prefers. Examples are questions like "what's your opinion about..?", free association of words, completion of sentences, and completion of narrations;
- closed questions: the client is asked to provide an answer inside a range of pre-defined answers. The most used closed questions are:
    - dichotomies: questions with only two possibilities of answers (yes/no);
    - multiple choice: questions with more than two possibilities of answers;

    - Likert scales: sentence to which the client has to provide his level of agreement;
    - semantic differential: the client has to choose among a scale of expressions between two opposite ones;
    - importance scales: the client is asked to evaluate the importance of specific features;
    - scoring scales: the client judges some features from the best to the worst one;
    - acquisition propensity scales: the client has to estimate his acquisition propensity;
- filter questions: they allow the interviewer to pass entire sections of the questionnaire depending on how the client answers to them;
- check questions: they verify the reliability of some answers previously provided.

In order to make the interview as effective as possible, some useful indications can be given [20]:

- follow the free flow of the speech: if the client is providing interesting information, it is not important to follow a prefixed path in the interview but it is recommended letting the client say what he wants to say;
- use visual supports, such as slides: some preliminary concepts of the product can be showed to the client to obtain a first reaction to the various solutions;
- avoid misguided hypotheses: the interviewers should avoid leading the discussion with hypotheses about specific project functionalities. When clients refer to technologies or characteristics of the product, the interviewer should investigate the latent need that the client is freely bringing out;
- ask pros and cons of the current available tools: in this way it is possible to bring out new client's needs simply by analyzing the constrains and benefits of the current solutions;

- pay attention to non verbal expressions: some needs can not be communicated by words, especially the one that involve interactions between man and machine, like graphics, style and simplicity. It is important to pay attention to the client's non verbal messages, like face expressions in particular moments of the interview;
- don't give for granted the meaning of words: sometimes users refer to their past experience and take for granted a particular meaning that they assign to a word. Sometimes this meaning is not the same as the common meaning assigned to that word;
- don't orient the answer with the question: questions must be impartial and neutral;
- try to involve only a subject inside the question: it is important to focus only on one subject at a time;
- avoid too long questionnaires: it is recommendable to limit the list of questions only to the fundamental ones.

2.4.3 Identify and organize needs

Clients, expressing their needs, will formulate free thoughts in an unstructured and raw way. So it is necessary to interpret these requirements in order to structure them and to create an organized list of what the needs to satisfy are. Some useful indications can be given in order to formulate the user needs in a correct and effective way:

- express the needs in terms of what the product must do, not in terms of what it could do: the need must be formulated independently from a particular technological solution;
- express the needs at the same detail level of the raw information, in order to avoid losing its original meaning;
- use positive instead of negative expressions, because it is easier to translate them in a product specification;

- express the needs through product attributes: it guarantees the coherence of the result and facilitates the translation in product characteristics;
- avoid using terms like “should” and “must”: these terms indicate a level of importance of the needs, which will be decided in a more structured way in the next step of the process.

Not all the needs emerged in this step have the same level of importance. A useful tool for classifying the requirements is the Kano matrix, as shown below:

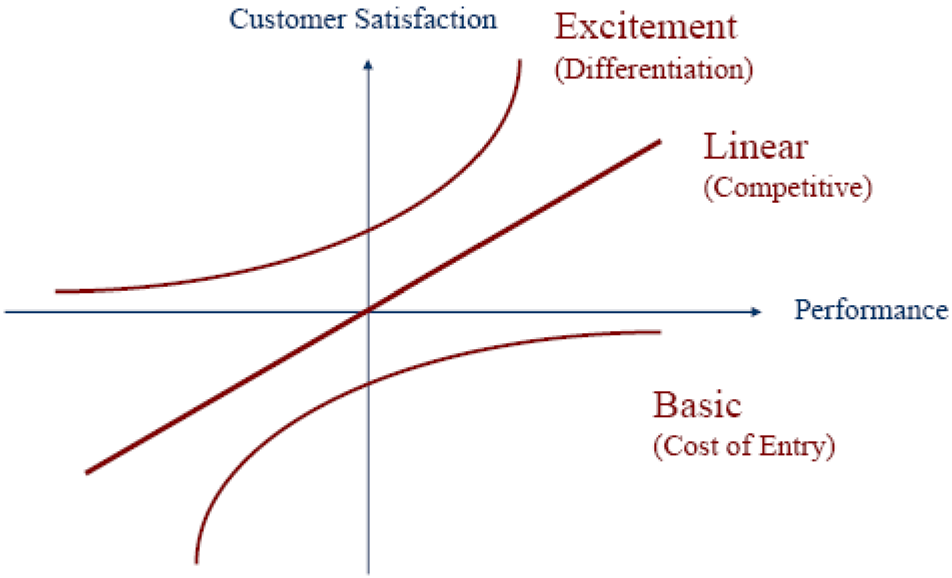


Figure 2-9 - Kano matrix

Following this representation, user needs can be divided in three different classes:

- excitement: requirements that the client doesn’t expect to find. So their lack will not be noticed, while their presence will make the user really satisfied;
- linear: requirements towards which the client expect an average performance but that, if performance is better than expected, it will results in a satisfaction increase depending on the distance from the average;

- basic: it doesn't matter how well this kind of requirement is satisfied, simply the user doesn't accept its lack thereof.

Excitement needs will have a higher level of importance, since they can greatly satisfy the user without an excessive effort. Linear needs, instead, will have a medium level of importance because an effort in this kind of requirement will be accordingly recognized by the client. Basic needs, at last, are not worth much effort because only their presence must be assured.

These kind of classification can be derived from several market researches, or directly asking to the customer through question such as:

- what are the main important needs among the ones that have been brought out?
- would you choose a service instead of another only because it satisfy this need (referring to one in specific)?
- would you be minded to additionally pay for a service that satisfies this need? How much is its value added?
- would you ever implement a tool not satisfying this need?

After having classified the needs in one of these three types, it will be easier to attribute a numerical index of importance to each requirement, depending on the client satisfaction it is able to guarantee.

2.4.4 Define product functionalities

The basic scope of this last phase is to translate user's needs in specific functionalities to implement in the final application. It is of utter importance to interpret the wishes of each and every customer into some tangible values that can be turned into engineering specifications.

Throughout the whole process, it can be useful to create and continuously update a list of the functionalities the can be implemented in the final application. After having identifying all possible users' needs, a list of possible functionalities of the final service has to be defined taking into consideration all that technology can provide. This is made through an in depth competence acquisition about the technology base and through a strict collaboration with the researchers that are physically developing the algorithms used for the final application.

After completing this list and making sure that no functionality has been left out, the entire set of functionalities can be analyzed through the two following dimensions:

- needs that the functionality satisfies;
- functionality feasibility.

After having assigned a numerical weight to both needs importance and functionalities feasibility, it is possible to create a mathematical model that gives a numerical importance value to each functionality taking into consideration both aspects.

2.4.5 Select the functionalities to implement

The selection of the functionalities to implement in the final application can be done in a structured way with the support of the House of Quality tool, as shown in the figure below:

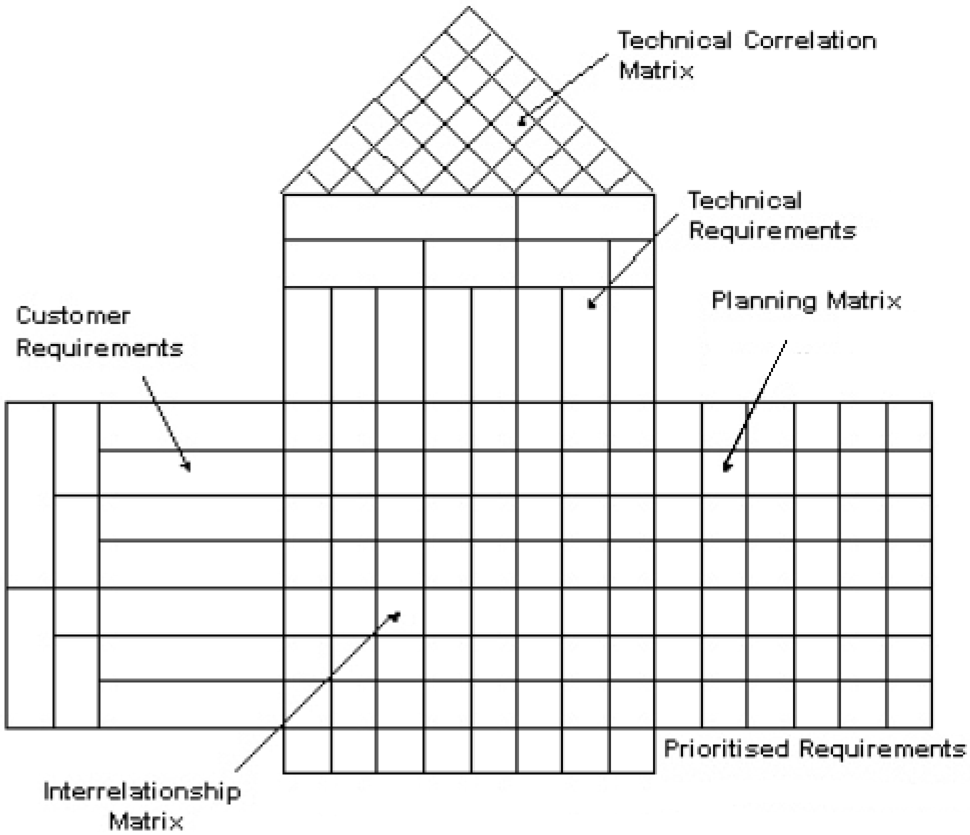


Figure 2-10 - House of Quality

Let's see in detail how each part of this matrix has to be built [21]:

- voice of the customers: in the left side of the House, customers' needs have to be determined, clarified, and specified. It is what has been done in the previous steps of the concept definition process. Now, all identified needed have to be organized in a list with a numerical weight assigned to each of them;
- technical requirements: the next step of the process is identifying which are the functionalities that can be implemented in the final service in order to satisfy customer's requirements. Once all needs are identified it is important to answer what must be done to the service design to fulfill the necessary requirements;
- planning matrix: the next step in the process is forming a planning

matrix, whose main purpose is to compare how well the team met the customer requirements compared to its competitors. Customer ratings are given to each company under each requirement. The customer ratings are combined with the weighted performance of each demand to produce an overall performance measure for the companies;

- interrelationship matrix: the main function of the interrelationship matrix is to establish a connection between the customer's requirements and the functionalities designed in the final application. It is necessary to assign relationships between each customer requirements and each functionality by symbols indicating a strong relationship, a medium relationship, or a weak relationship. The symbols can also be turned into respective indexes such as 9-3-1, 4-2-1, or 5-3-1. When no relationship is evident between a pair a zero value is always assigned. The interrelationship matrix should follow the Pareto Principle keeping in mind that designing to the critical 20% will satisfy 80% of the customer desires: therefore, there should not be a significant number of strong relationships between pairs. Having a look to the planning matrix placed on the left side of the interrelationship matrix, in addition, the company can begin to formulate a strategy to improve their service considering the strengths and weaknesses of the company and to determine what aspects need to be changed to surpass the competition, what aspects need to change to equal the competition, and what aspects will be left unchanged;
- technical correlation matrix: functionalities in existing designs often conflict with each other. The technical correlation matrix, which is more often referred to as the Roof, is used to identify where these units must work together otherwise they will be in a design conflict. The following symbols are used to represent what type of impact each requirement has on the other:
  - □ - Strong positive

- ™ - Positive
- x - Negative
- xx - Strong negative

These symbols are then entered into the cells where a correlation has been identified. Any cell identified with a high correlation is a strong signal to the team, and especially to the engineers, that significant communication and coordination are a must if any changes are going to be made. Many technical requirements are related to each other so working to improve one may help a related requirement and a positive or beneficial effect can result. On the other hand, working to improve one requirement may negatively affect a related requirement. If there is a negative or strongly negative impact between requirements, the design must be compromised unless the negative impact can be designed out. Negative impacts can also represent constraints, which may be bi-directional. As a result, improving one of them may actually cause a negative impact to the other. Sometimes an identified change impairs so many others that it is just simply better to leave it alone. Asking the following question when working with this part of the House of Quality helps to clarify the relationships among requirements: "if technical requirement X is improved, will it help or hinder technical requirement Z?";

- prioritised requirements: after having assigned a numerical value to both needs' importance and interrelationship between needs and functionalities, it is possible to create a mathematical model that gives a numerical importance value to each functionality taking into consideration both aspects.

At least, in this way, it is evaluated how each functionality could satisfy each need if implemented in the final version of the application and it is finally possible to define which functionalities to implement in the final service and which not to.

3 Overall on the European GeoPKDD project

The purpose of this part of the work is to shed some light on the GeoPKDD project: its objectives, its scope and background, and an in depth analysis of the project up to this point. The project will also be presented from a technological point of view, highlighting the current technological constraints and the research fields on which the project must work. The role of this work inside the overall GeoPKDD project will also be contextualized.

3.1 Objectives of the project

GeoPKDD it's an European project that aims to provide knowledge analyzing people movement through data coming from their mobile devices. A first overview on its objectives can be summarized by the acronyms:

- **Geo**graphic: the project deals with movement of people in space;
- **P**rivacy-aware: the project wants to preserve people's privacy;
- **K**nowledge: the project aims to find information about how people move in an aggregated way;
- **D**iscovery: the project had to develop algorithms to analyse movement data;
- **D**elivery: the project has also to package this information in useful services.

In our modern day society, information is ever growing and more easily obtainable than ever before. With this in mind, the GeoPKDD project is born; the idea of this project is to take all the data available through all the communication devices that use the telecommunications technologies of today and tomorrow (GSM, UMTS, GPS, Galileo…) and try to find a useful way to use this information, in order to improve the society and people's way of life. In reality, though, the possibilities of using this information are endless, ranging from communication network optimization, to improving traffic management, all the way to more commercial applications, such as geo-marketing and so on.

The main problem which must be considered is the users' privacy. That means that it must not be possible to find out sensitive information about any single person from the databases, and so all the analyses must be in aggregated form; therefore what the project is looking for is quality information regarding groups of people which can then be elaborated on to help in the decision making process.

The whole project's objective is to develop theory, techniques, and systems for knowledge discovery and delivery. Every time an entity moves through space, it creates a trajectory representing the history of its past and current locations. Examples of interesting trajectories of moving entities are hurricanes and tornadoes, animals and planes. Now, since GeoPKDD is dealing with a trajectory, the data it is taking in account is measured both in space and time: this means that the so called spatio temporal data must be defined in four dimensions. The fact that many of the techniques necessary for analyzing this kind of data don't even exist today must be taken into account, and therefore they will have to be developed during the duration of the project.

In fact, up until now the applications which have been developed are location-based services (LBS), which can be re-put as "give me some service depending on where I am now"; examples of this include applications in tourism, marketing, and transportation management. What the project wants to offer are movement-based services (MBS), where the new core information will be the cumulative movement of people, or in this case, the trajectories of the device users; the new service can then be rephrased as "give me some service depending on where I and other people have been in the past".

Therefore, it is necessary to go beyond the collection of positioning data sets to the delivery of information and knowledge derived from all the data that it's now possible to collect, and a framework which can serve as a road map for developing a privacy preserving knowledge discovery process must be identified. This brings to the three key objectives of the consortium:

- to devise methods for representing, storing and managing moving objects, which change their position in time, and possibly also their shape or other features, together with their trajectories, with varying levels of accuracy and certainty;
- to devise spatio temporal knowledge discovery and data mining methods and algorithms for moving objects and their trajectories;
- to devise native techniques to make such methods and algorithms intrinsically privacy preserving, as data sources typically contain personal location aware sensitive data.

Since the project is dealing with spatio temporal data, adequate information storage systems must be devised to allow easy access and retrieval. Technologically speaking, the idea is to try to take existing technologies which are today applied to a simpler version of information, such as location-based info, and build on them so as to reach the project's goals. This means creating a data warehouse to store objects and their trajectories over time, and then develop particular data mining algorithms and query languages which will guarantee the preservation of privacy for the end users.

3.2 The GeoPKDD process

The GeoPKDD process has been structured in 5 phases, also known as work packages; there are three main activities which make up the core process, and two orthogonal activities which will give support to the main tasks. The activities are not meant to be sequential, since much work has to be done in each one, but what's important is interfacing the different teams working on different parts of the project in order to maximize the synergies between them. The three main research areas include the privacy-aware trajectory warehouse, the privacy-aware spatio-temporal data mining methods, and the geographic knowledge interpretation and delivery. The work in each of these research parts will be brought along by the other two activities: application demonstrators, and harmonization and dissemination.

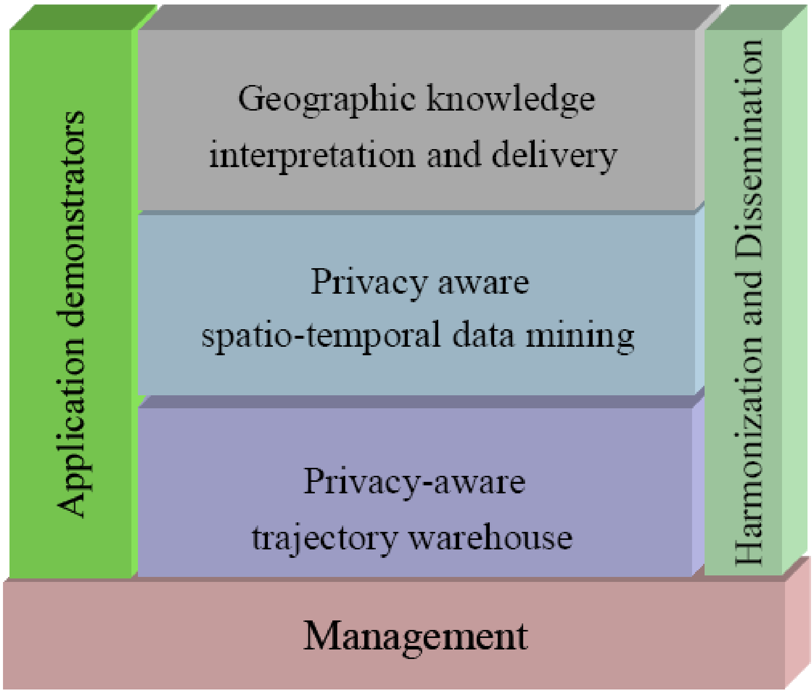


Figure 3-1 - The five Work Packages

Let's explore each one of the activities more specifically and which are the results of the GeoPKDD project in these fields up to now [27].

3.2.1 Privacy-aware trajectory warehouse

The goal of this activity is to devise a model for building a trajectory warehouse, that is a data warehouse specialized for the management and analysis of a large collection of trajectory data. The model and the architecture will be used as basis for the activities in the other work packages. Another important goal is to investigate privacy issues related to a data warehouse containing trajectory data, while developing the model and the architecture. Therefore, the goal is to develop a data warehouse system that has all the built-in mechanisms required to ensure privacy and security.

Let's briefly define what a data warehouse is: a data warehouse is the main repository of an organization's historical data, its corporate memory. It contains the raw material for management's decision support system. A data warehouse differs from a database because it is non-volatile, meaning that data in the data warehouse is never over-written or deleted: once committed, the data is static, read-only, but retained for future reporting.

The main problems of the trajectory data warehouse involve the multi-semantic spatio temporal data it has to store; indeed, the main sources of complexity relating to spatio temporal data are:

- the huge size of data, which calls for efficient access methods based on ad-hoc indexing schemes, data compression techniques, etc.;
- a very large searching space, which requires the use of background knowledge and powerful heuristics;
- the continuous streams in which data come, requiring the ability to incrementally mining the data stream in an efficient and accurate way.

A trajectory data warehouse allows the user to execute a series of steps:

- transform location data to trajectories, allowing the user to tweak the streaming parameters according to the type of data which is analyzed;
- store trajectories;
- query trajectories, through predefined or customized queries;
- enable the extraction, transformation and loading process to fill the data cube;
- provide external links to call a local OLAP viewer that offers multidimensional analysis and allows users to quickly analyze a great quantity of complex data.

3.2.2 Privacy-aware spatio temporal data mining methods

The goal of this work package is to design, implement and experiment algorithms for spatio temporal data mining, specifically meant to extract spatio temporal patterns from trajectories stored in the data warehouse.

Data mining is a complex iterative and interactive process, which involves many different tasks that can bring the analyst from raw dirty data to actionable knowledge. A rigorous user interaction must take place in order to facilitate efficient knowledge manipulation and discovery. Such a rigorous interaction can be achieved by means of a Data Mining Query Language (DMQL), which should include:

- the specification of the source data to be mined;
- the kind of knowledge to be mined;
- background knowledge and symbolic reasoning on it, in order to exploit such knowledge to drive the mining process towards those patterns that are more likely to be interesting;
- the definition of constraints that the extracted knowledge must satisfy;
- the representation of the extracted knowledge;
- the post-processing of the extracted knowledge, enabling the crossover between patterns and data.

In the mining field of spatio-temporal data, the connection between spatial and temporal attributes needs still to be clarified: it is already clear that time cannot be treated merely as an extra spatial dimension and that its linear and unidirectional nature, as opposed to the bidirectional nature of space dimension, has not been properly investigated.

3.2.3 Geographic knowledge interpretation and delivery

Extracted patterns are very seldom pre-packaged geographic knowledge; it is necessary to reason on patterns and on pertinent background knowledge, evaluate how useful the patterns are, overlay them with relevant geographic information, and suitably visualize these patterns.

Constructing a GKDD process through this perspective aims to integrate different reasoning tasks in a unified system by mapping the complex relationship between movement metaphors and patterns. This could be the most difficult part of the whole project, since this is an iterative process that is carried out by the experts of the application domain, and they are the people who input the qualitative part of the final output since they extract the hidden meanings from the pattern results of the data mining algorithms.

3.2.4 Application demonstrators

This activity involves a preliminary design of selected applications. Then, to demonstrate the GeoPKDD usefulness, special applications will be developed. They should be efficient, usable with a simplified GUI and tailored for the GeoPKDD demonstrator.

A great number of services which could be offered to a number of different clients can be visualized. The real objective of this activity, though, will be to provide the foundation of the MBS services, demonstrating with working prototypes the validity of the ideas underlying the project itself. In fact, the present day algorithms which are being used can only explain the patterns of the past trajectories recorded; what could be interesting, and should be the objective of future research, is to develop an updated version of the algorithm which will be able to give predictions on future trajectories of the users, since this kind of information will be infinitely more valuable in the field of new future applications.

3.2.5 Harmonization and dissemination

This work package should represent the interface with society, focusing on the public relations side of the project. The objectives of this activity include harmonizing the project with European and national privacy regulations and authorities, analyzing the potential applications and markets for the GeoPKDD paradigm, and defining the exploitation plan and dissemination of the project results in different forums.

It is of paramount importance to underline the preservation of privacy for every citizen belonging to society, especially from the point of view of the public opinion, otherwise it will be extremely difficult to take full advantage of the project's benefits; what could be seen is a "big brother", privacy-violation effect, where people will feel as though their personal rights as individuals will be violated. In light of this, the members of this activity will have to address the public relations part of the project as well, stressing the importance of privacy in all the work packages.

3.3 Analysis of GeoPKDD process

In order to underline the main features of the GeoPKDD project, it can be seen as a black-box process like in the figure below:

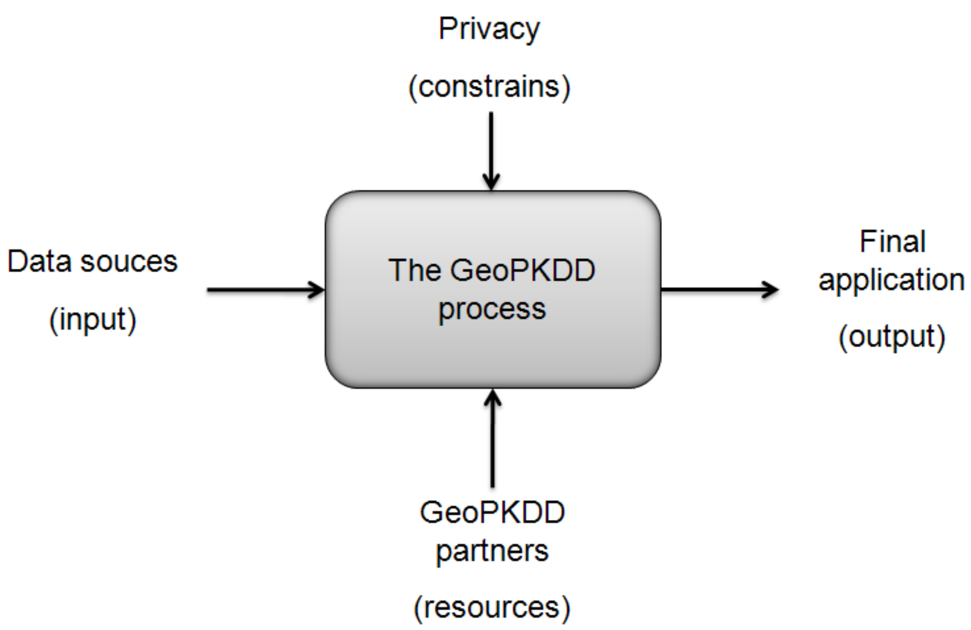


Figure 3-2 - Visualization of GeoPKDD process

3.3.1 Input

The input data for the GeoPKDD process comes from the mobile devices of the final user. Two positioning technologies are nowadays used by cellular telecommunication companies for mobile devices: GSM (Global System for Mobile Communications) and UMTS (Universal Mobile Telecommunications System) protocols. These protocols are based on a network architecture in which a geographical area is covered by a number of antennas, called BTS, which stands for Base Transceiver Stations; these antennas emit a signal which is received by mobile devices. Each antenna covers an area called "cell" which is uniquely identified by an id. In urban areas, cells are close to each other and the area covered is relatively small; the cell diameter can go from 100 meters up to a km; in rural areas, instead, the radius of a cell can reach a maximum of 30 km [26]. The presence of a device inside a cell is detected periodically by the system; all of the acquired data is then maintained in GSM and UMTS location databases called VLR and HLR (Visitor and Home Location Register):

- the VLR stores information about changes in position that occur between substructures of the network (between single antennas or groups of antennas) of the same location area, which includes a group of antennas covering a quite extended area (100-300 km^2), controlled by the VLR;
- the HLR, instead, stores information about changes of position between different location areas, and therefore between different VLR. These last kinds of registers maintain location information within substructures of the network such as location areas; the actual location of a device in the GSM network is also registered at this level.

However, since these registers are specialized in routing rather than positioning, the advent of Location Based Services in the last few years has brought along another class of location registers. These are called SMLC (Serving

Mobile Location Center) servers, and they can calculate the location of a local device through antenna power level comparison between different cells.

Location raw data can be expressed in different data formats, depending on which method is used to calculate the position on the network. In general, the data format of the location information should indicate:

- the time of detection, that can be represented by a tuple indicating the detection date and time;
- a position information represented by a precise location with some geographic reference system;
- the identifier of the detected mobile user.

The most used cellular position techniques to calculate an approximate position of a mobile device are:

- CI (Cell Identity): in this method, the position of a device is identified by the cell where it is connected. This information is available at the network as well as at the handset. The antenna identifier is converted to a geographic position by means of the existing knowledge residing in the coverage database SMLC. Accuracy depends on the cell size and the antenna type and can vary from 100 m to few km.
- Cell Identity + Timing Advance: this method improves CI using measurement reports that contain power level at the handset from the serving cell and cells on the neighbour list. The power level at the handset can be used to estimate distances from device to antenna. Accuracy is slightly better than CI, and depends on the location of antennas and environmental conditions that can affect the signal strength.
- E-OTD (Enhanced Observed Time Differences): in this positioning method, the handset measures the arrival time of signals transmitted from 3 or more antennas. Two specific methods can be implemented: E-OTD MS-Assisted, in which measurements are made by the handset and then transferred to the SMLC that

calculates the position of the device, and E-OTD MS-Based, in which the position calculation function resides at the handset. Accuracy can vary from 50 to 100 m. It is worth noticing that, in terms of resources, E-OTD is a very expensive method since it needs some additional and specific equipments to be added to the network.

In addition, there are other positioning methods based on another technology, the GPS (Global Positioning System) protocol. GPS is a global system of geographical location which exploits a constellation of satellites allowing the location of a mobile device with an accuracy of 50-80 m. Examples of positioning methods using GPS technology is A-GPS (Assisted GPS), in which the handset measures the arrival time of signals transmitted from 3 or more satellites. This technology has a quite low impact on the network because it requires only the support at the SMLC level. Positioning performances are better in rural space and poor in urban space where buildings and other obstacles disturb signals from satellites. This method is quite efficient, in terms of quality/cost ratio, and reliable in terms of quality of given information. On the other side, it has a technological constraint, since the handset needs to be GPS compliant to receive signals from the satellite.

The collection of such a huge amount of stream data represents the basic data source of the GeoPKDD process; however, the availability of such data is not always immediate: the location registers are critical databases for the owner company, mainly due to privacy concerns, making difficult the access to log data. The data availability problem is the main reason why many data generators have been built. Generators are tools that, given some input parameters and an output format, generate synthetic datasets in order to help in testing algorithms over different scenarios on datasets with predictable characteristics.

As new technologies have been producing new types of raw data (spatio temporal data), research has focused on creating theories and methods for understanding and processing these new kinds of data. A new standard for this data has been created as an evolution of existing data types, through the addition of a temporal dimension. An algebra of operators has been defined to provide full facilities for expressing queries involving moving objects.

3.3.2 Inside the process

The term “knowledge discovery in databases” was coined in 1989 to describe the overall process of identifying valid, novel, potentially useful, and ultimately understandable patterns in raw data coming from the data sources described earlier. The knowledge discovery process can be divided in the following nine steps:

- developing an understanding of the knowledge domain, the relevant prior knowledge, and the goals of the user;
- creating a target data set, selecting a data set, or focusing on a sub-set of variables or data samples, on which discovery is to be performed;
- data cleaning and pre-processing;
- perform data mining algorithms;
- evaluating and validating the results;
- consolidating discovered knowledge, incorporating this knowledge into the performance of the system, or simply documenting it and reporting it to users.

Although this list might suggest a sequence of steps, a knowledge discovery process is in fact a random process in which the steps are often carried out by an unsystematic approach and do not follow a straightforward analysis.

Moreover, these steps have been treated as separate activities, with their own principles, procedures and limitations, making necessary an effort in integrating the heterogeneous results.

Knowledge in GeoPKDD must come from viewing the patterns and integrating them with the use of metaphors. Indeed, it is only after the metaphor makes sense that an unknown set of patterns can be interpreted and understood by the expert of the application domain. Metaphors are artifacts of knowledge which can specify a concept from a domain in terms of another. In the GeoPKDD context, metaphors help the comprehension of what makes a pattern structurally and meaningfully different from another. Ideally, metaphors should be built in the expert domain, since they require a high level of abstraction which could make sense within a specific context; then, and only then, it will be possible to extract useful knowledge. Metaphors are not directly correlated to reality, but they are useful if they can help structuring it.

The reasoning can be defined as the ability of the expert to formulate and operate with concepts in abstraction, such as metaphors. The logic of discovery can be obtained through three different approaches, depending on the kind of reasoning made [27]:

- deduction: the reasoning task through which it is possible to infer a consequence from a set of patterns. These consequences are inferred from general (patterns) to specific (metaphors). The metaphor, in this case, is already known by the experts and usually represents the empirical base of the GeoPKDD process, since the relationship between metaphor and pattern can be verified directly;
- induction: the reasoning task through which it is possible to infer a generalization from a set of patterns. This implies reasoning from detailed facts (examples) to general principles (conclusions). This is the approach of learning from examples, where the example is the metaphor which contains more information than the ones contained in patterns. An example is the “movement as activity” metaphor, which expresses a set of activities requiring the existence of a person in a specific place at a certain temporal sale (daily, weekly, monthly). If the activity is used to explain the extracted pattern, then a certain set of “generalized” activities must be previously known;
- abduction: the reasoning task through which it is possible to infer the best cause for the occurrence of a set of patterns. It is quite flexible since it is not constrained to the use of pre-existent metaphors, but it is instead free to create new metaphors that can help to explain the patterns. In spite of this, data mining methods don’t work this way, but rather by deduction or induction.

The discovery of knowledge is not a banal process and it requires the examination of metaphors in order to understand through patterns what the world is like. A way to describe the process of knowledge discovery in the GeoPKDD process is the use of ontological tiers, which establish the movement metaphors and integrate different reasoning tasks in a unified system. Each tier uses the metaphors instantiated on the previous tier, enabling the understanding of interesting and meaningful patterns. The tiers used in this prospective are:

- tier 0: the reality space. This tier recognizes the existence of a known world made of four dimensional fields: three in space and one in time. The expert has to formalize background knowledge in order to generate an a-priori knowledge made of defined hypotheses and explicative rules;
- tier 1: the positioning space. Observations are values computed at each point of space and time, working on a value scale which could be either quantitative or qualitative. In this level, only a set of observations of spatio-temporal localizations is available, represented by points in space: a representation of the trajectory of these points is not available yet. It is nevertheless possible to distinguish observations according to four types of points:
    - stop: a set of points representing stop of a very short time, like a few minutes, due to traffic lights or stop signs;
    - stop over: a set of points representing a change of speed, such as a road accident;
    - short stay: a set of points representing stays of short time, that is some hours, due to activities such as work, shopping or leisure;
    - long stay: a set of points representing stays of several hours, such as devices left at home or being turned off.

  The metaphor mostly used in this tier is the one of movement as urban shape: since human behaviour is constrained to urban morphology, such as transportation network or density of a city, this metaphor is necessary in order to infer knowledge from the point generated at this level;

- tier 2: geographical space. This tier represents the geographical space where the cognitive system is able to extract trajectories. The movement of an entity is defined as a trajectory, which is a set of lines between stops, stop-overs, short stays and/or long stays. The metaphor which is mostly used at this level is the one of accessibility, that can be obtained through the calculation of distances. It is inferred using the average distance between centroides of the leaving zone and the destination zone. This kind of computation is not suitable to the network transportation in order to establish the real distances of the routes: as a matter of fact, they are not based on direct lines between origin and destination, and they don't consider the real trend of urban morphology. Therefore, the use of the "trajectory as urban shape" metaphor becomes necessary, already used in the previous tier, together with the one of accessibility, in order to generalize trajectories using the transportation network;
- tier 3: the social space. This tier includes the model which merges daily trajectories to the fundamental human activities. At this level, geographic environment is considered as a space where human activities take place. A social system is formed by people, groups and organizations which maintain relationships through intentional activities based on rules and values. So it is important to realize that a trajectory takes place not only in space, but in a socio-spatial system. The metaphor mostly used in this tier is the one of movement as activity. The socio and spatial systems are strictly related to each other and they can not be analyzed in a separate way. Processes in the social system, such as economy, politics and culture, have consequences in the spatial system, and vice versa. The typical reasoning task used at this level is induction and the goal is to identify the best model that can approximate trajectories; for

example, it is possible to find:

  - patterns that can explain the occurrence of a certain activity, such as shopping or leisure;
  - dependencies between different characteristics of activities;
  - sub-groups of activities and time intervals with similar patterns;
  - occurrence of an unexpected activity;

- tier 4: the cognitive space. The goal of this tier is to obtain knowledge by means of adductive reasoning tasks which can work also without pre-defined hypotheses or rules. Information has to be integrated with a set of personal knowledge of the experts, so that it can be opportunely interpreted. It is important to underline the difference between tacit knowledge and implicit knowledge. The latter is something that the experts should know but don't want to express, while the former is something the experts know but they can not express: it is something personal, difficult to communicate, and it is not easily expressible through the formality of language.

3.3.3 Constraints

Privacy is a very delicate issue which has to be addressed in the whole GeoPKDD project, from the beginning of the geographic knowledge discovery process up to the integration between patterns and metaphors.

The concept of privacy is usually transferred into the one of protection of personal data, which can be defined as any information relative to an identified or identifiable person. This definition focalizes the attention on data as opposed to people and considers the possible identification of a person as the real problem. Thus, an activity is privacy-aware if and only if it doesn't involve the identification of a person, through a set of attributes.

Anyways, literature concerning privacy suggests a new vision of privacy, which defines the person as a definition of profiles independently of the individual identification. As a matter of fact, since trajectories are integrated with an a-priori knowledge about places and social contexts, data miners can infer the individual identity through the identification of profiles, or through the classification of people into pre-defined groups. Although these groups are built from aggregated information which are not related to a single person, they could cause the discrimination of groups. In this way, all members in a specific area or trajectory can be categorized in one group, and the identified category could be attributed to all members.

In addition, positioning data coupled with geographic information allows many inferences in the knowledge of the social background. As a matter of fact, the number of interpretations of being in a place at a certain time is often limited, for example the group of people who goes to visit the same doctor who is specialized in AIDS treatment. Another reason is that there are some spatio-temporal models that are related to a pre-determined human behaviour, for example the place where a person spends the night without moving is almost surely his home.

Furthermore, it could be possible to reconstruct the individual identity of the subject even when explicit identifiers, such as name or phone number, have been erased or substituted with pseudonymous. As a matter of fact, there are some attributes that, in combination with others, can bring to a “re-identification” of people. A set of attributes which can be uniquely related to individual identity is called “quasi-identifier”.

As seen, positioning data are strongly identifying and, since users can not avoid being in a place at a certain time, their impression of lacking control on surveillance is particularly felt (the big brother effect).

There are three fundamental approaches that can be used in order to guarantee the privacy preserving:

- knowledge hiding: this approach, well known as “sanitation” of data, aims to hide specific sensible knowledge; this can be obtained modifying data in the database;

- data hiding: this approach, well known as “distribution reconstruction”, acts only on the data source in order to avoid the identification of the original raw database, while at the same time it allow the reconstruction of the distribution of data at an aggregated level;
- distributed data hiding: this approach, well known as “secure multiparty computation”, is based on cryptographic techniques and aims to create a data mining model from different distributed datasets, where each part doesn’t communicate its own data source to the other parts. Each user, therefore, can have access only to the determined set of data for which he is enabled by his own identifying key.

The GeoPKDD project moves towards a new emerging research focusing on a potential privacy violation of extracted patterns [26]:

- secure knowledge sharing: while the analyzed models focus their attention in the privacy preservation of the model, this technique focuses on the privacy preserving of the data mining results. Instead of sharing the data, it shares the association rules (that is the output of the model) and imposes that a certain set of association rules must not be disclosed; so this technique requires a classification of association rules based on their perceived privacy and their ethic sensitivity. The “sanitation” is made at association rules level, and not at raw data level.

In order to correctly address privacy issues, the project has to take into consideration the real protagonists of the process, that are stakeholders. Each stakeholder has different levels of interests relative to GeoPKDD. Three main stakeholders can be identified:

- users, that is the ones who generate positioning datasets. They should authorize the level of privacy requested, the use, the openness and the individual participation;
- data collectors, that is the ones who are interested in collecting

positioning data and in developing data mining algorithm. They should assure the requested level of privacy, the quality of data and the security safeguards;

- experts of the application domain, that is the ones interested in the application of the output of GeoPKDD process. They should define the level of privacy requested for the collection, the scope and the results use.

Data itself is neither good nor bad, but it is its use and scope that brings the distinction between acceptable or unacceptable. In order to characterize how trajectories are used and from whom, the following metaphors can be introduced:

- the Spy: the one who has the goal of finding knowledge about the behaviours of people in order to investigate or supervise. He searches for traces in the present and usually, but not always, he is motivated by bad intentions;
- the Historian: the one who aims to make archaeological investigations. It means that he wants to characterize the behaviour of the community in order understand its dynamics and the way in which it lives. He searches for traces in the past and usually, his intentions are good.

Nowadays the project is trying to provide the archaeology of present times: this means that the goal is to find in real time precise representations of what people do through the digital traces of their mobile activities. Therefore, the project is trying to make it possible to learn from recent history in order to design efficient services that can improve the way people live today. This knowledge of traces, from a few moments ago or perhaps yesterday, can improve the activities of decision making relatively to mobility issues.

A third subject in these metaphors can be introduced:

- the Scientist: the one who aims to prevent the Spy from violating individual privacy and data anonymity, without preventing the Historian from conducting his studies.

The problem is to find an optimum trade-off between two opposite goals:

utility and privacy. While respecting utility, the project has to provide useful knowledge to the Historian and, at the same time, to deny access to sensible information to the Spy. Any real solution to the problem can be obtained only through a partnership between technology, legal issues and social rules: it is with this goal in mind that the observatory of privacy has been created within the GeoPKDD project.

3.3.4 Resources

According to the project work plan, a project Consortium of eight members all around Europe has been established, with careful balance of needed complementary expertise: limited replication of competences has been sought in fundamental issues, notably spatio temporal privacy-preserving data mining, where research is at a very embryonic stage and tighter collaborative work is needed.

Here is a table summing up who the eight participants belonging to the consortium for the GeoPKDD project are [27]:

Partic N.	Participant name	Short name	Country
1	KDD Lab, joint research group of ISTI-CNR, Istituto di Scienza e Tecnologie dell'Informazione, Pisa and Univ. Pisa Dept. of Computer Science	KDDLAB	Italy
2	University of Limburg, Theoretical Computer Science Group	LUC	Belgium

3	EPFL Ecole Politechnique Federale de Lausanne e University of Milan	EPFL	Switzerland
4	Fraunhofer Institute for Autonomous Intelligent Systems	FAIS	Germany
5	Wageningen UR, Centre for GeoInformation	WUR	Netherlands
6	Research Academic Computer technology Institute, Research and Development Division	CTI	Greece
7	Sabanci University, faculty of Engineering and natural Sciences	UNISAB	Turkey
8	Wind Telecomunicazioni SpA, direzioni reti Wind progetti finanziati & Technology Scouting	WIND	Italy

Table 3-1 Participant list

The project will be coordinated by KDDLAB; the scientific coordinator will be Fosca Giannotti. Of course, the project will need a critical mass of competences, due to the interdisciplinary nature of the topics addressed; let's now define the distinguishing skills and responsibilities of each participant:

- KDDLAB will mainly contribute in WP2 to the investigation of privacy-preserving spatio temporal data mining algorithms, in particular clustering and frequent patterns; it will also contribute in WP3 to the design of a data mining query language, and in WP5 to the activation of an observatory on privacy regulations. KDDLAB is the project coordinator, and more specifically coordinator of WP2 and WP5;

- LUC will mainly contribute in WP2 to algorithms for spatio temporal association rule mining and in WP3 to the design of pattern languages; it will also contribute in WP1 to the trajectory data warehouse;
- EPFL will mainly contribute in WP1 to the specification of the trajectory data model and uncertainty management, taking into account multiple resolution and multiple representations of ST data, as well as security issues;
- FAIS will mainly contribute in WP3 to the geographic visualization of spatio temporal patterns; it will also contribute in WP2 to the development of spatio temporal mining algorithms and shall have a key role in the dissemination action. FAIS will coordinate work package 3;
- WUR will mainly contribute in WP4 to the application of the various components for the GeoPKDD process and to the selection and design of a societal application. WUR will coordinate work package 4;
- CTI will mainly contribute in WP1 to the identification of privacy threats for location-based data and trajectory data and to techniques to address such threats, especially in the trajectory warehouse but also in the OLAP and data mining algorithms; it will also contribute in WP2 to the development of methods for privacy preserving data mining. CTI will be the coordinator of WP1;
- UNISAB will mainly contribute in WP2 to the investigation of privacy-preserving data mining techniques;
- WIND, being a telecom operator, will play its main role in WP4 in providing input raw data and background knowledge about mobile clients and cellular dimensioning, and in the deployment of the GeoPKDD prototype in a demonstrator for mobile network

optimization.

3.3.5 Output

The output of the GeoPKDD process is the delivery of the extracted knowledge by means of useful services for the final user.

Until now, the most diffused services based on localization data are real time services depending only on where the user actually is, and do not take into account any historical information. In order to guess what kind of services could be built from positioning data, let’s take a look at the main kind of location based services that nowadays market offers [28][29][29][30][31][32][33]:

Service	***Description***	***Features***	***Kind of business***	***Business potentiality***	***Type of service***	***Availability***
Location-based push services	Services that allow subscriber people to get real-time information based on their location and their preferences, such as personal advertising	• Periodic localization • User profiling	B2C	High	LBS	Yes
Location-based pull services	Services that answer to particular needs based on user's localization, such as general information (weather, traffic etc) or individuation of the nearest bar, pub, restaurant, bank etc.	• No user profiling • Localization on demand	B2C	High	LBS	Yes

Emergency services	Services that can emit automatic alarms to the public force basing on where the user is	• Pull service • No commercial aims	B2C	Low	LBS	Yes
Assistance services	Services that allow people to have assistance and reparations basing on their localization	• Pull service • Occasional use	B2C	Medium	LBS	Yes
Assets Tracking	Services that help tracking a fleet of mobile assets (cars, camions, busses etc.) for identification, assistance, safety, optimization or logistics. Typical users are transportation, car rental, insurance or logistic companies	• Push service • Suitable only with GPS devices	B2B	High	LBS	Yes

Community services and games	Services that keep people in contact with their families and community, thanks to a location based chat or games	• The user is localized by a third person • The use is for personal aims	B2C	Medium	LBS	Yes
In-vehicle guidance	Services that supply people with the best route to reach the final destination and with other useful information such as traffic and eventual re-routing in case of variations	• Push service	B2C	Medium	LBS	Yes
Pedestrian guidance	Services that help people to reach the final destination by foot and supply them with useful information about the surrounding environment	• Pull service	B2C	Low	LBS	Yes

Car sharing	Service of sustainable mobility that permits a more efficient use of cars within the city. Thanks to a GPS device on the mean, it is possible to take its moves under control	• Push service • Need of a user login and account	B2C	Low	LBS	Yes
Info parking	Service that communicates to the final user the availability of parking spaces in the neighbourhood	• Pull service	B2C	Medium	LBS	Yes
Hello bus	Service that permits to know in real time the remaining time from the next arrival of a bus to a particular stop, by means of SMS or call center.	• Push service	B2C	Medium	LBS	Yes

LocHNESs by Telecom Italia	Service that allow making previsions about traffic analyzing users' trajectories. Data are obtained by the position of users' mobile devices	• Push service • No user's profiling since raw data are in anonymous format • Suitable with both GSM and GPS data	B2C or B2B, depending on final use	High	MBS	Not already commerciali zed, only tested in experimental way
Connected Car project, by TomTom and Vodafone Netherlands	Solution that can convert raw GSM data into real-time information on the speed and direction of cars traveling throughout the road network	• Push service • No user's profiling since raw data are in anonymous format • Suitable with GSM data	B2C or B2B, depending on final use	High	MBS	Not already commerciali zed, only tested and proven on a regional scale

Table 3-2 - LBS and MBS services

As it is possible to guess from the table, positioning data are mainly used in order to offer LBS services, with the only exceptions of few MBS research projects, still at an embryonic state. The market is full of services offering real time information basing on the present positioning of the user: the main goal of the GeoPKDD project, instead, is to help citizens in their mobile activities by analyzing the traces of their past activities. This analysis is made through data mining methods.

The definition of possible outputs is the main task in which this work is involved. The basic idea is that each important technology innovation must be accompanied by a market research in order to understand how these new technologies can be packaged and delivered to the final user. Without the conception of possible usable services for the end consumer that can actually function, no technology can really be effective in improving life quality and in raising the innovation level of society.

3.4 Where this work stands

After this in-depth look at the main features of the GeoPKDD project, let's try to understand which part this work plays inside the whole project.

WIND, being the only commercial partner of the GeoPKDD Consortium, is really interested in understanding which could be the market potentials of this project and how it will be possible to create a gainful business from it.

This brings to the final goal of this work, that is to try to focus on "marketing" oriented services and to have a deep look at their potential market. The basic idea is to understand how these technical researches can be used to build up a final service that can represent a potential business opportunity for WIND and that can justify its involvement in the GeoPKDD project. Different kinds of clients will value different services; the important thing is to understand what kind of service will be most appreciated by each, taking in consideration what the output of the data mining algorithms currently available is.

It's of paramount importance to underline that this work must focus on producing a result that will be in line with the kind of data that are at disposal. This means that, even though a whole series of intelligent and useful services could theoretically be developed, they couldn't be implemented as of now, since they would require much more accurate information from the users, or more elaborate data mining algorithms that could give future predictions on users' movement.

Therefore, what this work is supposed to provide to the GeoPKDD project is:

- a scouting of the possible business services that can be implemented taking advantage of the developed technology;
- evaluation of these services in order to individuate, among the list, which could be the best service to carry out in the development process;
- in depth analysis of the service's market and of the client's needs, in order to understand what kind of requirements the user expects from the final service;
- service definition, along with the specification of the functionalities that need to be implemented.

Next step of the project will be the development of a prototyping demonstrator of the selected service in order to test its feasibility and to show to the interested users the potentialities of the work done so far.

Concluding, what this work leaves to the GeoPKDD project is the definition of a final service that needs to be implemented and that can lead the several European partners' work to a unique vision. As a matter of fact, up to now all work packages have been carried out separately, with a lack of integration. This was caused by the fact that there was not a clear idea of what kind of concrete application these researches would be turned into. That's the scope of this work and its main contribute: a clear direction for future integration among several researches, and an unique final vision of what has to be actually implemented and sold from the GeoPKDD project.

4 Concept generation for the GeoPKDD project

4.1 Generation of ideas

In order to allow a creative generation of ideas about which kinds of services could be brought out from the technology at disposal inside the GeoPKDD project, a brainstorming session was carried out on September 21st, 2007.

The purpose of the brainstorming was to give a panoramic view of the possible final services that could be brought out from the GeoPKDD project. The participants of the meeting were senior consultants and team leaders of CONSEL Academy and they shared their cumulative experience gained in years of projects in the consulting field.

Participants Name	Participants Competence
Gennaro Galdo	Senior Consultant
Francesco Limone	Project Manager
Marco Amici	Senior Consultant
Luigi de Costanzo	Project Manager
Filippo Conci	Senior Consultant
Gianluca Sabatini	Senior Consultant
Matteo Pieragostini	Team Leader

Table 3 - Participants to brainstorming session

In the brainstorming process a lot of services were brought out.

In order to structure the process, the ideas coming from all the participants were divided into three main classes, according to what they referred to:

Figure 4-1 - Three groups of ideas

In this way, it was easier for the members to organize their own ideas and to create a good reasoning pattern that led to the definition of a final service.

The complete list of clients, needs and services that were brought out in the brainstorming meeting is shown, together with the logical connections between them.

Legenda:

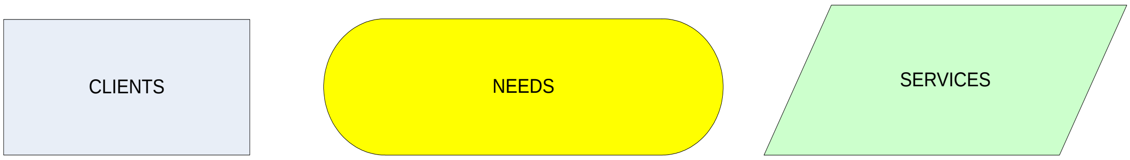


Figure 4-2 - Legenda

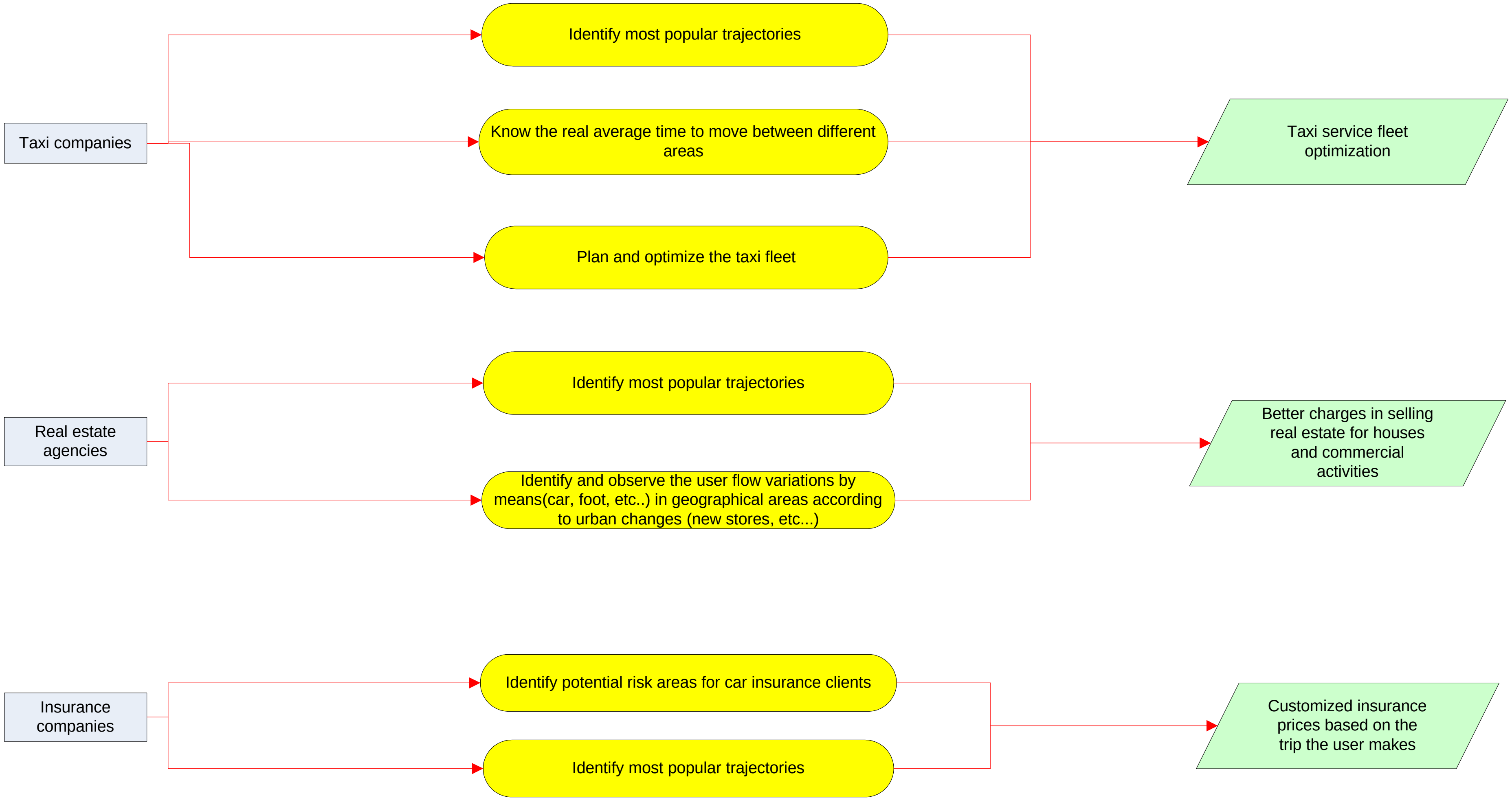


Figure 4-3 - Brainstorming output part 1

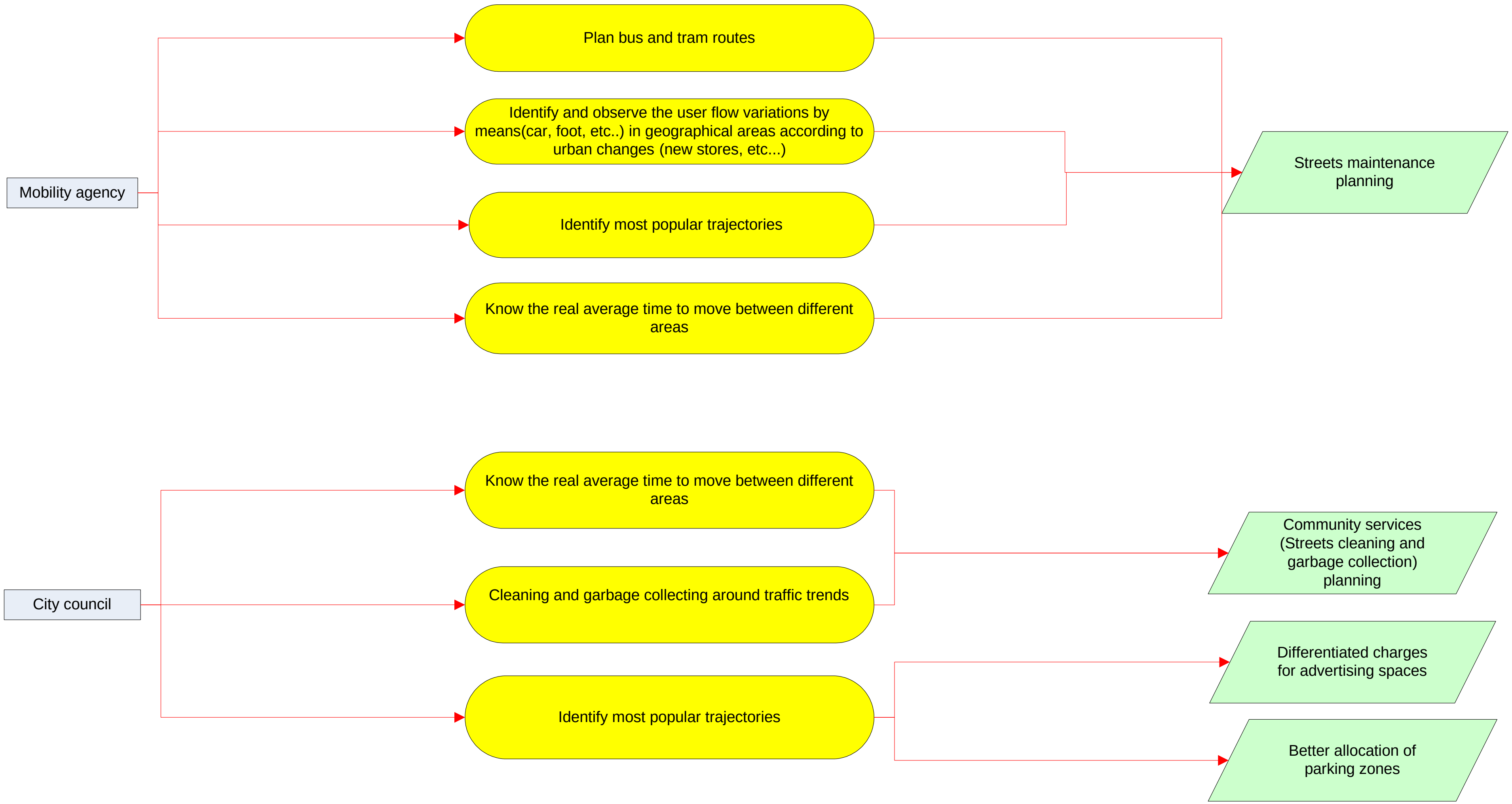


Figure 4-4 - Brainstorming output part 2

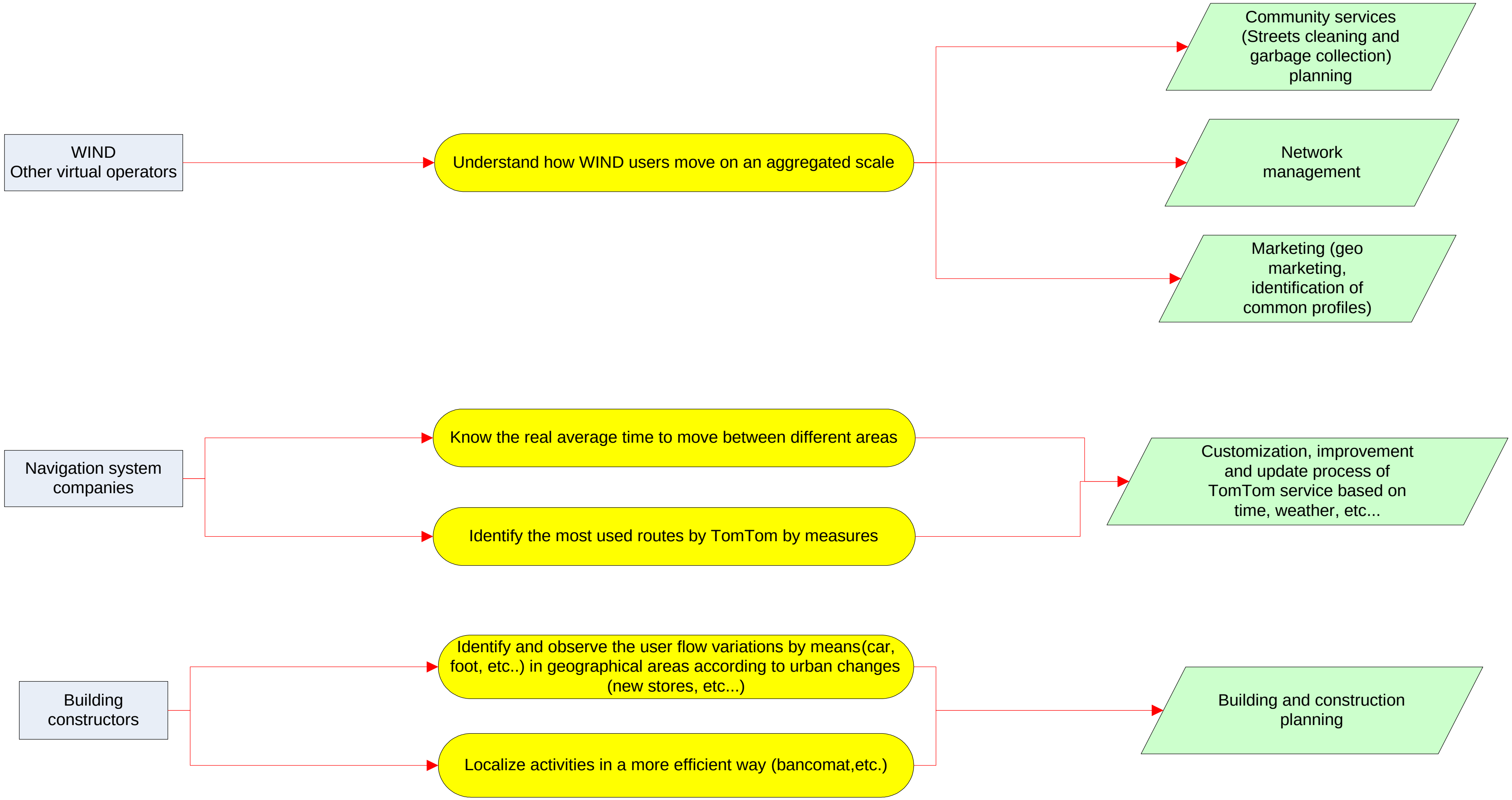


Figure 4-5 - Brainstorming output part 3

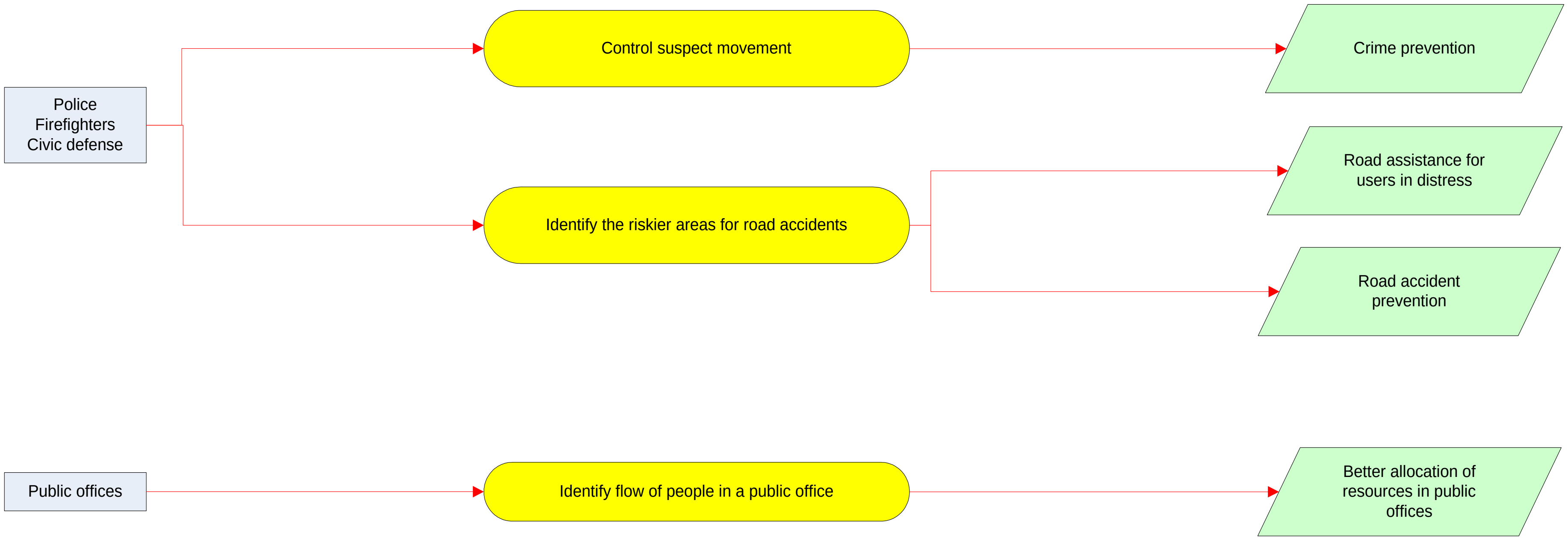


Figure 4-6 - Brainstorming output part 4

4.2 Concept screening

Raw data coming from the brainstorming session was managed in order to complete the list of concepts with new and interesting ideas. First of all, all clients nominated in the brainstorming session were taken into account in order to bring out new needs that could characterize each of them. Then, the list of needs was analyzed in order to generate new unexploited services that could satisfy these needs. This last complete list of services was finally taken as the real input of the further steps of the whole process.

After the generation phase, a first screening of the several service concepts was made in order to focus the further efforts only on the most interesting ones.

These are the tiers of the screening phase:

- elimination of all pointless ideas;
- grouping together of all similar concepts;
- evaluation of each idea following these two criteria:
  - simplicity of implementation;
  - possibility of creating a profitable business;
- selection of the first 7 ideas.

The screening process was performed asking the participants of the brainstorming session to evaluate the several ideas with reference to one medium concept. As reference concept, the geo marketing support tool was chosen by the majority of the participants, since it was considered to have medium features with respect to other ideas. The results of the screening process are shown in the figure below:

Service Name	Simplicity	Profitability	Total
Marketing support tools (geo marketing, differentiated charges for advertising spaces, identification of common profiles, etc)	0	0	0
Customization, improvement and update process of TomTom service based on time, weather and other events	+	+	2
Service for customized insurance prices based on the trips the user makes	+	+	2
Traffic management and urban planning tool	+	+	2
Public means management tool	+	+	2
Network management service: bandwidth allocation, antennas positioning, dynamic power assignment and so on	+	+	2
Final services for tourists and statistics about their behaviours (for example optimizing trips based on the time at disposal or providing SIM cards with some free hours of calls in order to observe their movements)	+	0	1
Strategic localization planning of opening of new activities	-	0	-1
Buiding and infrastructure location planning and pricing	-	0	-1
Tool that defines better charges in selling real-estate for houses and commercial activities	-	0	-1
Travel packages for tourists based on common patterns	0	-	-1
Creation and update of virtual maps of the more frequented places, locals, shops and so on	0	-	-1
Common profile identification of users of a telco based on their behaviors for entertainment services (chat, games, and so on)	+	--	-1

Figure 4-7 - Screening process part 1

Service Name	Simplicity	Profitability	Total
Taxi service fleet optimization	-	0	-1
Identification of riskiest areas for prevention of road accidents	--	-	-3
Support for a more efficient allocation of resources in public offices	-	--	-3
Street cleaning planning	--	-	-3
Garbage removal trajectory and frequency planning	--	-	-3
Tourist information points allocation	--	-	-3
Planning of assistance for users in distress in presence of extraordinary events	--	-	-3
Trashcans optimal allocation	--	-	-3
Instrument to be used for psychological studies on different classes of people (children, elderly, prisoners)	--	--	-4
Streets maintenance planning	--	--	-4
Crime prevention support tool	--	--	-4

Figure 4-8 - Screening process part 2

In this document the focus is only on the first seven services, explaining their relevance to the GeoPKDD project, the user needs they satisfy, and the type of users they address.

Geo-marketing support tool

This concept was chosen as the reference for the evaluation of the other concepts since, for all the participants of the brainstorming meeting, it has medium simplicity of implementation and possibility of creating a profitable service out of it. The service can be an important tool in order to support the decisions about how the advertising spaces around the city can be priced in a differentiated way, and how to better distribute the advertising materials or to locate new commercial activities.

With a geo-marketing service supported by the information of GeoPKDD application, it could be possible to satisfy several types of needs: to localize activities in a more efficient way, to identify and observe the user flow variations by means (car, foot, etc..) in geographical areas according to urban changes (new stores, etc...), and to create mutual deals between activities belonging to the same flow (geographical area).

The possible final users of this kind of service are mainly advertising agencies, marketing companies and business intelligence companies that, through this data, can extract useful information about the actual movement of people in the urban network depending on the period of the day or the day of the week. It could be possible, for example, to better profile citizens and to understand each market segment's behavior, in order to better allocate advertisement efforts and to maximize the return on investments in marketing campaigns. It is worth to underline, anyway, that in this field privacy issues have to be carefully taken into account.

Telecommunication companies tool

The idea for this service was brought out starting from the analysis of the previous concept. The basic idea is the following: if WIND has data about its users' movement, why can't WIND itself use this information? As a matter of fact, knowledge about clients' movements can be a precious information in order to better profile users and to better invest in retention campaigns that, nowadays, are an important investment field for telecommunication companies. Using this information it could be possible, for example, to decide where to put billboard for retention campaigns or to select the best place to open a new store. The main advantage of this idea is that it could be feasible to create a close loop model in which WIND is not only the data provider but also the client itself, passing in this way all the problems related to privacy issue.

The needs that this service tries to satisfy are to identify the most popular trajectories and to understand how WIND users move on an aggregated scale.

The immediate clients of this kind of service could be WIND and also other telecommunication companies and other virtual operators such as Carrefour, Coop, Poste Italiane that are going to enter the mobile phone market relaying on an external network; as a matter of fact, these last kinds of clients could use the information about their own users' movements in order to better allocate advertising or shops positions.

GeoPKDD project can also be used in providing information about the best allocation of bandwidth for the single antennas that make up the network itself. It could also provide useful information on the optimized location of the antennas in a city in order to better cover phone and network traffic and satisfy the needs of the customer. As a matter of fact, understanding the way phone users move on an aggregated scale is an important a must-have for any telecommunication company.

Customization, improvement and update process of TomTom service based on time, weather, etc.

This application refers to an update on the already existing service developed by TomTom and other similar navigation systems companies. The new features will predict the real flow of cars at each time and make the route planning of GPS devices more effective based on the knowledge of the past trajectories of people during the day. This kind of knowledge can help making forecasts about the traffic flow according to weather conditions or the happening of some special events like concerts, football games, strikes, and other social events. Taking into account this information, the application can calculate which are the best routes.

This service was selected as being the most promising among all services, because it obtained the highest evaluation from all the members that participated in the screening process.

The application addresses navigation system companies like TomTom and satisfies the needs to identify the most used routes by people in cars and to know the real average time to move between different areas. The input data for this kind of service would need a very high accuracy, because it would be necessary to understand the actual road taken by the users, and therefore GPS data would seem like the best fit.

Traffic management and urban planning tool

This can describe a service for mobility agencies. The service will help the agency to forecast traffic based on different periods of time and also in the presence of extraordinary events (football matches, concerts, etc…), to better plan the future maintenance of the roads, and to allocate the public transportation resources in order to satisfy the travel requirements of their customers. It could have a humongous impact on urban network, for example making a street one way or two ways, changing the timing of the traffic lights or the traffic signs.

Another functionality which could be very important is allowing the user to simulate future traffic trends in presence of modifications in the road network; there are already many software tools that allow to do this, but they are all based on theoretical models that can only estimate what the traffic flow is, while through historical analysis, the GeoPKDD algorithms could give a much more precise idea of how traffic could adapt to future modifications of the roads.

The input data for this kind of service can be very diverse, and according to the type of data at disposal, different functionalities can be implemented. In this case, both GSM and GPS data can be used, along with data already at disposal coming from traffic cameras and traffic sensors.

The needs that led to this kind of service were: identifying and observing the user flow variations by means (car, foot, etc..) in geographical areas according to urban changes (new stores, etc...), identifying most popular trajectories and knowing the real average time to move between different areas.

Public means management tool

This service could be very useful for public transportation companies all around the world, since analyzing the routes of public means along with historical data coming from private mobility could facilitate the client in optimization of routes and the distribution of bus and tram stops. Along with these macro-functionalities, using GPS data coming from the means themselves, it could also be possible to understand the behaviors of bus and tram drivers, and an automatic driver reporting tool could also be developed with ease.

Many public transportation companies today have mounted GPS devices on all their means, and through these it is possible to obtain a measurement of the position of each mean every thirty seconds, with a very accurate measurement of speed, course, and lots of other derived indicators. Another important information that busses can provide is the number of people inside the vehicle at any time, through motion detectors mounted on each of its doors. It could also be possible to integrate this precise data about busses with GSM data coming from private vehicles, in order to develop an accurate traffic flow model.

Customized insurance prices based on the trip the user makes

The participants proposed to use the GeoPKDD information in order to create a new pricing model for the insurance companies based on the identification of the riskiest roads that a user normally takes. With a GPS device the companies can track where the user base and the single driver moved and, based on that information, customize a different offer for the insurance price. It would also be possible to understand the behavior of single users at the wheel, for example if he is a speeder, or very cautious, how soon he breaks at stops, if he takes the highway most of the time he is driving, etc…

This service would satisfy he needs to identify potential risk areas for car insurance clients and the most popular trajectories, but also allow an insurance company to better calibrate their risk models through a more accurate calculation of the policy price for each customer.

The addressed users for such a service are insurance companies.

Similar solutions are already being developed in Italy by OCTOTELEMATICS, a company who installs GPS devices on board of cars, and then analyzes the data in order to understand the behaviour of its users. What is lacking though, is a historical analysis of the cumulative data, and this is where the GeoPKDD algorithms could provide added value.

Final services for tourist and statistics about their behavior

Knowing the time that it takes for a tourist to travel from one point to another point, a service can be brought out. Through a Personal Digital Assistant (PDA), It will be possible to optimize the tourist trips and to generate customized routes which can be changed at any time based on the time they have at disposal: for example, it would be possible to have a route planner installed on the PDA, that could have as an input for the itinerary the Points of interests that the user wants to reach, along with the amount of time at disposal; it would give multi-modal routes (car, public means, foot), and it would also store in its memory multimedia data that would be available in the vicinity of the POI, along with useful information (for a museum, it could show ticket price, opening and closing hours, along with the most famous works of art inside it).

This service would also be useful to make statistics on the most visited places, allowing different pricing policies for different touristic areas, and also for obtaining a more accurate calculation of the flow estimate of tourist behavior, which could be useful both for hotels, which could plan more interesting offers for their customers, and also for stores, bars, and restaurants, which could work out a deal in order to provide offers via LBS services to tourists once they are in the proximity of a certain point of interest and their store.

To improve the service all tourists can receive a mobile device that could be either a phone and a SIM card so that they can have the opportunity of making some free calls in a national area making the service more attractive, but also allowing GeoPKDD to observe their movements and infer information from these data, or a GPS phone in order to make the localization more accurate.

These services would satisfy the needs to identify the most visited spots by tourists and to know the real average time to move between different areas.

The final users would be tourists, tourist operators, and hotels.

4.3 Concept scoring and selection

Now that the number of ideas is reduced, each of them has to be evaluated in order to create a final and structured score of them.

In order to score all the different concepts that have been described so far, it was necessary to find an appropriate tool that could, given a definite number of drivers, put all the services on the same plane, and allow a fair confrontation in order to understand which of them would be worth investing in.

In this case, the chosen measurement tool is the bubble diagram. This type of diagram takes into account three drivers (dimensions) on which the different concepts can be given a quantitative or qualitative score. While the first and second drivers correspond to the x and y axis, the third driver is represented through the width of the “bubble”, which represents the graphed item. In this way, it is possible to have a clearly understandable three dimensional analysis on a two dimensional plane.

In order to score the MBS services that came out of the brainstorming, three fundamental drivers were identified: business appeal, technological feasibility, and WIND data consistency. While the first two were already taken into account during the brainstorming process, the third dimension was the added value for this particular analysis, since it is a key issue that needs to be considered in order to maximize the client’s needs. An important thing that must be mentioned is that for the scope of this work, it was chosen to classify the services on a qualitative basis, which means that each service was defined on a scale of LOW, MEDIUM, HIGH, for each driver. A more specific definition of the three drivers will now be given:

- technological feasibility: the issues that this particular driver wanted to capture include low initial investment, and therefore an easy access to the potential market in terms of the technology needed to set up the product line, the training of personnel, etc..; potential compatibility with the existing infrastructure which implies the possibility of utilizing the technology which the company already owns; ease of implementation: this would include specific training of new personnel, possible new costs deriving from this service, like costly maintenance procedures, an eventual transitory phase needed to set up new equipment, and so on. This driver represents the third dimension in the bubble diagram, and is defined through colors, such that since each service will be represented with a circle in the final bubble diagram, a red circle will mean low technological feasibility, yellow will stand for medium technological feasibility, and green will be high technological feasibility;

- WIND data consistency: this is the key driver for the client company. As the reader will already know, the developed technology can work with both GPS and GSM data; depending on the kind of data at disposal, different accuracies can be obtained, and different kinds of services can be though of: WIND, as a telecommunications company, has at disposal the data coming from its GSM network, which has the advantage of having at hand data for the whole country, and the disadvantage that its accuracy is definitely raw. Keeping this in mind, the two issues that this driver wanted to capture were how the service fits data at disposal, and the availability of needed data; deciding to implement a service specifically thought for GPS data would result in no added value to the final customer, which would in turn mean investing in something that would be sure to fail. In the bubble diagram, WIND data consistency will be measured on the y axis;
- business appeal: this is the most important driver out of the three, and it gauges the business potential of the new service that is being measured through a number of key issues: potential market, which includes the number of potential customers that could subscribe the service, the potential segmentation of the market, and so on; profit

on the single service, which includes the amount of mark up profit that is possible to make on each instance of the service: in a mass product market, for example, competition is very fierce, and so the margins that a company makes on the single sale are very low: this implies concentrating more on mass market penetration than anything else; service innovation, which means understanding how different the product that the company is launching on the market is from anything else that is already available, since having a clear idea of what is available on the market is very important information in order to define the marketing strategy; potential competitors, which, as before, implies understanding what other companies are present on the market and with which products, but also forecasting how these competitors could respond in the future. This key driver is represented on the x axis, and can assume low, medium, and high values.

Now that the three drivers have been identified, let's have a look at the final results of the scoring process:

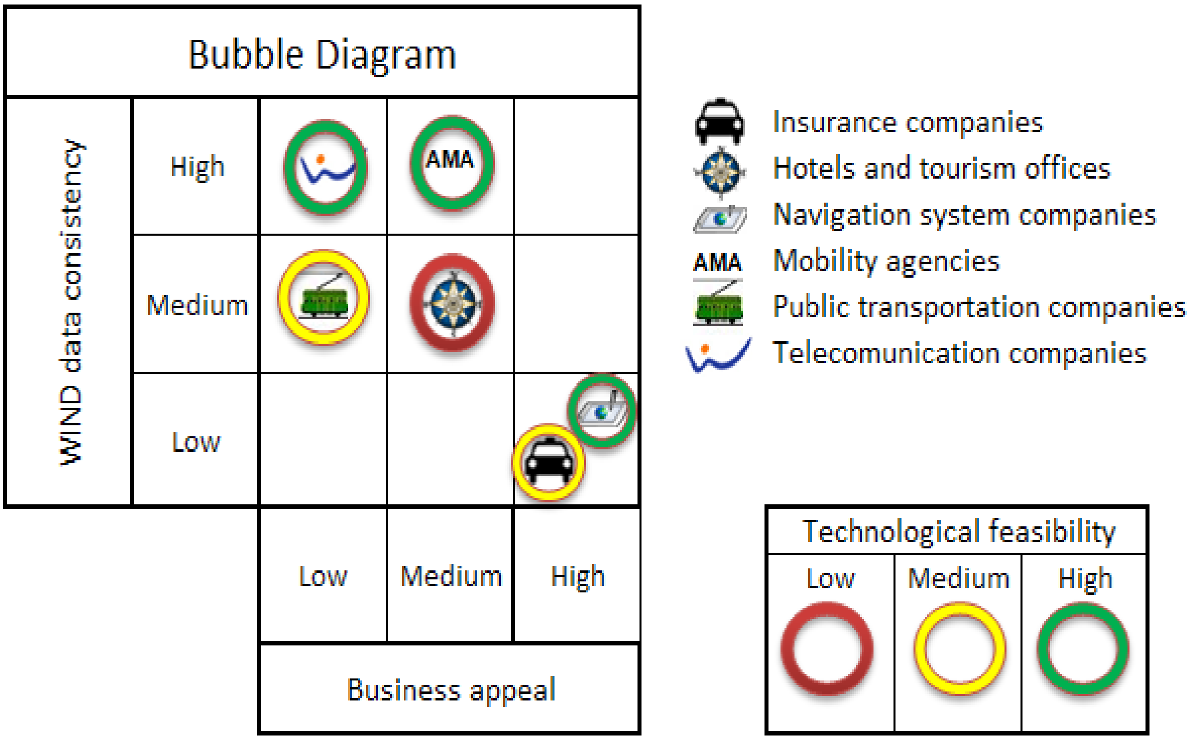


Figure 4-9 - The completed bubble diagram

These are the six services represented through the customers they address. Taking into account the three drivers shown previously, this is how they rank. It is now possible to choose two of the six and concentrate future energy on them. The two services that were picked are the one for mobility agencies (traffic management tool), and the one for WIND (geo-marketing and network planning tool).

5 Definition of functionalities for the GeoPKDD application

5.1 Initial market analysis

5.1.1 Market identification

The market in which the service enters is the one of the mobility agencies and traffic managers.

This market is completely new for the Wind, and new is also the kind of service that Wind is going to offer. Therefore, as it is possible to see from Ansoff matrix, the strategy required for entering the new market is diversification:

	Existing product	**New product**
Existing market	**Market penetration** Increase sales to the existing market Penetrate more deeply into the existing market	**Product development** New product developed for existing markets
New market	**Market development** Existing products sold to new markets	**Diversification** New products sold in new markets

Figure 5-1 - Ansoff matrix

In particular, Wind is going to undertake a conglomerate diversification, since it is going to move in new target markets with a service that is totally unrelated to its existing business.

The market of reference of this service involves those actors working in the management of public and private mobility, both at urban and regional level. More in detail, the target market for this kind of application is composed by the following realities:

- traffic and mobility management departments of local public administrations. This first part of the market is formed by 109 unities, that is the number of Province in Italy;
- traffic and mobility management departments at regional level. This second part of the market is formed by 20 unities.

It is important to underline that, among the 109 Province, some of them have already given birth to a specialized mobility agency, which can be either private or public and that is involved in traffic and urban mobility issues. Let's have a brief look about the role of mobility agencies and the reason why they were created.

With Dgls 422 of 99 and Dgls 400 of 99 (Art. 35) the market of public transportation was liberalized by the Italian Government, in order to create new companies which could be independent and with autonomous power over the company management, policies and incentives. In 2002 this was the situation of the market of public transportation companies:

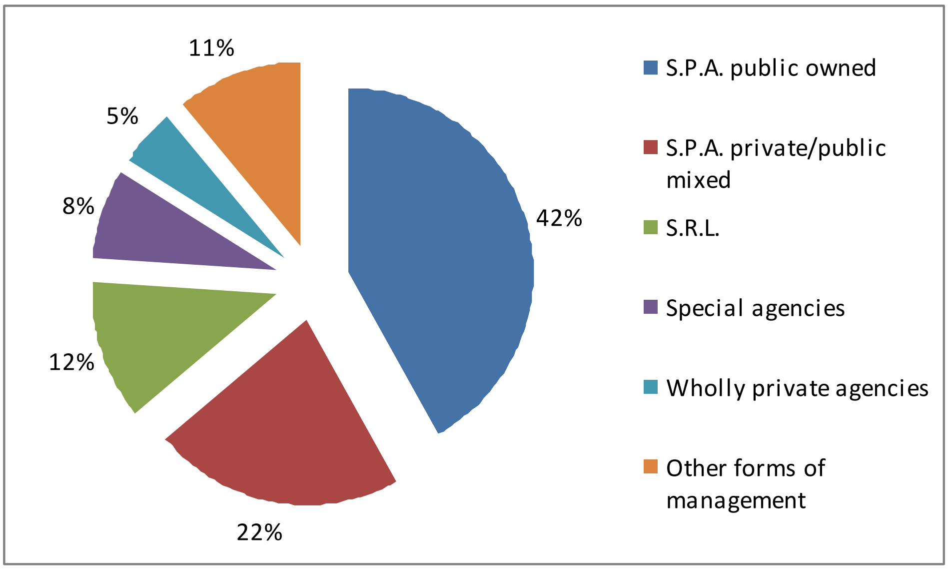


Figure 5-2 - Public transportation companies market analysis

This kind of reform brought an increase in efficiency and effectiveness of public transportation companies, with an higher level of investment in new technologies and in innovation than in the past. The market became competitive while before was monopolistic; in addition it assisted to separation of the ownership of infrastructure from its management and the separation of programming activities from management ones. The second type of separation brought to the creation of new intermediate organisms called Mobility Agencies with technical competences to which all planning decisions are delegated. Mobility agencies have the role of planning and designing the integration between public services and private mobility, along with the management of mobility demand, access to city centre, parking policies and other services.

In order to have a clear picture of the different tasks that different mobility agencies can carry out, the figure below shows a comparison between some of the companies in the market and their fields of interest:

Agencies	Agenzia ACT Reggio Emilia	Agenzia ATR Forlì/Cesena	Agenzia di Parma	Agenzia di Piacenza	Agenzia Milanese Mobilità e Ambiente	Agenzia per la mobilità metropolitana di Torino	Atac Roma	Agenzia mobilità Genovese
Legislative form	S.P.A (Provincia e Comuni)	S.P.A (Provincia e Comuni)	S.P.A (Provincia e Comuni)	Consorzio provinciale	S.R.L. (Socio unico metropolitana Milanese)	Consorzio provinciale	S.P.A. (Comune)	S.P.A. (Comune)
Activities of planning and design	yes	yes	yes	yes	yes	yes	yes	yes
Design and management of mobility services (scolastic, handicaped)	yes	yes	yes	yes	yes	yes	yes	yes
Design and management of complementary services (car sharing, road pricing)	yes	yes	yes	yes	yes	yes	yes	yes
Parking management	yes	yes	no	yes	yes	yes	yes	yes
Tariff income management	yes	yes	no	no	no	no	yes	no
Delegation and management of service contracts	yes	yes	yes	yes	yes	yes	yes	yes
Ownership of infrastructural assets	yes	yes	yes	no	no	no	yes	yes
Ownership of means	yes	yes	no	no	no	no	yes	no
Means maintenance	yes	yes	no	no	no	no	no	yes

Table 5-1- Functionalities of some mobility agencies

With reference to the possible functions, all mobility agencies carry out planning activities with a different level of integration and regulation. Other functions, even if they are carried out by some, can not be considered typical of a mobility agency. The regulation activities include among other things the calculation of prices, fares and grants, individuation of efficiency and quality standards, implementation of incentives or sanctions and so on.

5.1.2 SWOT analysis

Before furthering the analysis of the target market previously identified, it is worth to keep the focus more general and to analyze the market at a higher level in order to identify the main features that characterize it and to have a clear picture of the business in which this new service enters. This analysis was carried out from two different points of view:

- external, in order to highlight market's threats and opportunities;
- internal, in order to highlight Wind's points of weakness and strength.

The results of this SWOT analysis is proposed below:

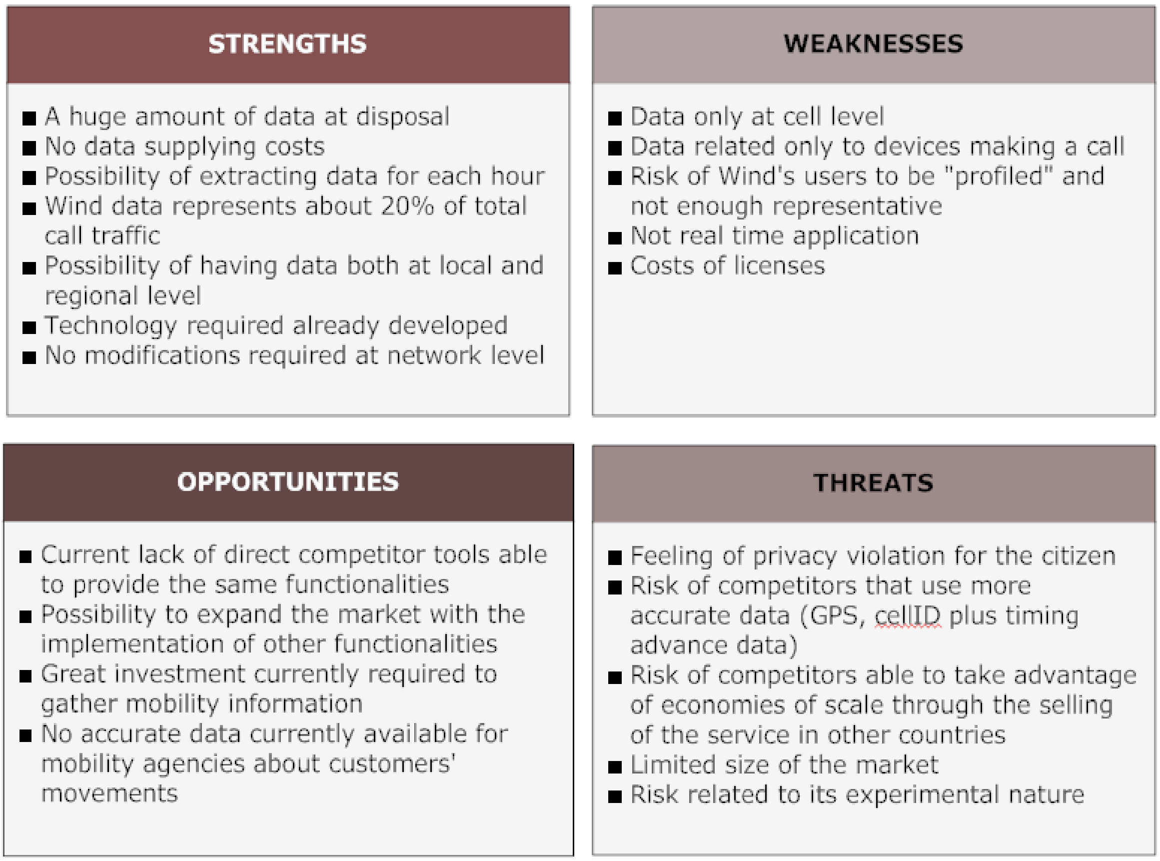


Figure 5-3 - SWOT analysis for WIND

From this external and internal analysis, some general considerations can be brought out about the service that WIND is going to provide.

The opportunities are really high, especially for the fact that no competitors are currently able to produce the same kind of data and that the availability of this kind of information is greatly required by the customers. WIND, being the first mover in the new market, could easily raise entry barriers so that other companies will have problems in becoming competitors. An example of entry barrier could be the creation of long term relationships with the clients, that can be carried out by the selling of the licence for a long period and by a good customer care service. Furthermore, the great interest showed by the market for this innovative service, especially by the “A” type segment of the market, makes it possible to gain a wide piece of the most attractive market in a short time, so that future competitors will have to invest a lot of efforts to gain the less attractive part of the market. Another positive feature that makes this business so appealing for WIND is the relative simplicity of carrying out the service: as a matter of fact, data is already at disposal and it has only to be downloaded, while the other tools required for storing and visualizing data have already been developed by GeoPKDD Consortium. Basically, the efforts required are really small in comparison to the profits that can be obtained.

The only risks that need to be underlined are related to the quality if data at disposal. As a matter of fact, GSM data provided by WIND has a low accuracy in comparison with GPS data. Anyway, it is also true that cell phones are so diffused that no other data can be as representative for the totality of citizens. Mobility agencies or, more in general, traffic managers are interested in studying the general trends of urban flow in order to make aggregated analyses about people movements. This means that the availability of a huge amount of data representative of the totality of urban movements is needed, and only GSM devices can assure such diffusion. Therefore, it can be stated that there are enough strengths inside WIND to face these external threats.

5.1.3 Market segmentation

The totality of the market is composed by 129 unities but, among them, some differences can be found.

The main discriminative criteria is the creation of a mobility agency for the urban traffic management. As a matter of fact, it is plausible to suppose that the Province having a mobility agency can be more attracted by the service that is going to be developed than the other ones because their field of interest is focused only on mobility issues and so they can be much more inclined to invest in ITS applications. In addition, through an initial market analysis it was found out that mobility agencies are investing much more financial efforts in innovative projects (like info-mobility projects and so on) than traffic departments of public administrations where a separate mobility agency was not creating.

Up to now, 13 mobility agencies were established and they can be considered as type “A” clients for the reasons explained above. Nevertheless, there are 5 regional laws that oblige Province in creating a mobility agency for urban traffic management. The Italian regions involved in this kind of regulation are:

- Emilia Romagna, with 9 Province,
- Liguria, with 4 Province;
- Piemonte, with 8 Province;
- Lombardia, with 12 Province;
- Campania, with 5 Province.

It is reasonable to suppose that in these 38 Province a mobility agency will be established in the next years, so they are going to become attracted clients in the next future. For these reasons, they can be considered type “B” clients. All remaining Province, that are 71, have been instead considered as type “C” clients.

Concerning regions, a basic hypothesis was made, which is that the five regions imposing the regulation mentioned above can be considered more attracted by the application because they are investing a lot of efforts in improving mobility and so they are more receptive to the innovation provided by this project. Therefore, traffic and mobility department of Emilia Romagna, Liguria, Piemonte, Lombardia and Campania are considered as “A” type clients, while the other regions are considered “B” clients. Another important hypothesis is that no Italian region can be considered “C” type client: as a matter of fact, the potentialities of mobility management at regional level with this tool are really high, since movements in that case can be analyzed at location area level and so also inactive devices data is at disposal.

The results of market segmentation can be represented in the figure below:

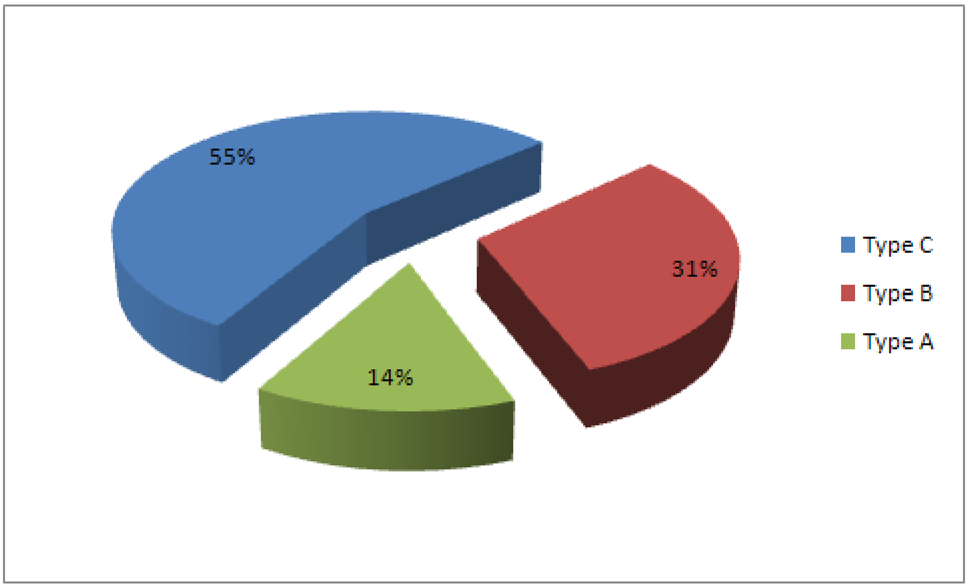


Figure 5-4 - Market segmentation

5.1.4 Demand forecast

Taking into account the previous segmentation, a method to estimate the future demand of the application was applied.

The proposed method evaluates future demand with this formula:

$$Q = N \times A \times P$$

where:

- N: number of potential clients;
- A: potential acquisitions;
- P: probability of acquisition for an informed user. It is possible to estimate P with this empiric formula:

  P = C surely × F surely + C probably × F probably

  where F surely is the ratio of people that stated that surely would buy the service; F probably is the ratio of people that stated a probable propensity to acquisition; C surely and C probably are calibration constants that usually assume the values 0,4 and 0,2 respectively.

Since the small size of the total market, it is reasonable to suppose that the market can be contacted in a complete way, so that A value is assumed to be 1. A sample of the three types of clients were contacted, in order to gather their acquisition propensity. Let's see the results obtained.

From all the 129 possible clients, 20 % of them were contacted, that is 25 clients. Among them, the trend of the acquisition propensity was the following one:

- sure willing of acquisition = 7 clients, that are mobility agencies of Ferrara, Rimini, Reggio Emilia, Forlì, Piacenza, Milano and Comune di Vercelli;
- medium acquisition propensity = 2 clients, that are mobility department of Emilia Romagna and the mobility agency of Roma, Atac.

That means that:

F surely = 0,28

F probably = 0,08

P = 0,13

Q = 17

It is worth to underline that 6 of the 7 clients who showed a great acquisition propensity belong to the type "A" clients, while the two clients who showed a medium willing of acquisition belong to the "B" type. This witnesses the validity of the market segmentation made above.

5.1.5 Industry identification

The tool that is the subject of this business model belongs to two different industries, each of which has its own features, competitors and trends.

As a matter of fact, this service can be considered as a MBS type of service (industry 1) that is applied to the ITS sector (industry 2). Both acronyms refer to a more general category of service which will be explained in depth in the following part of the document. The specific industry in which the service can be placed is the intersection between them, as shown in the figure below:

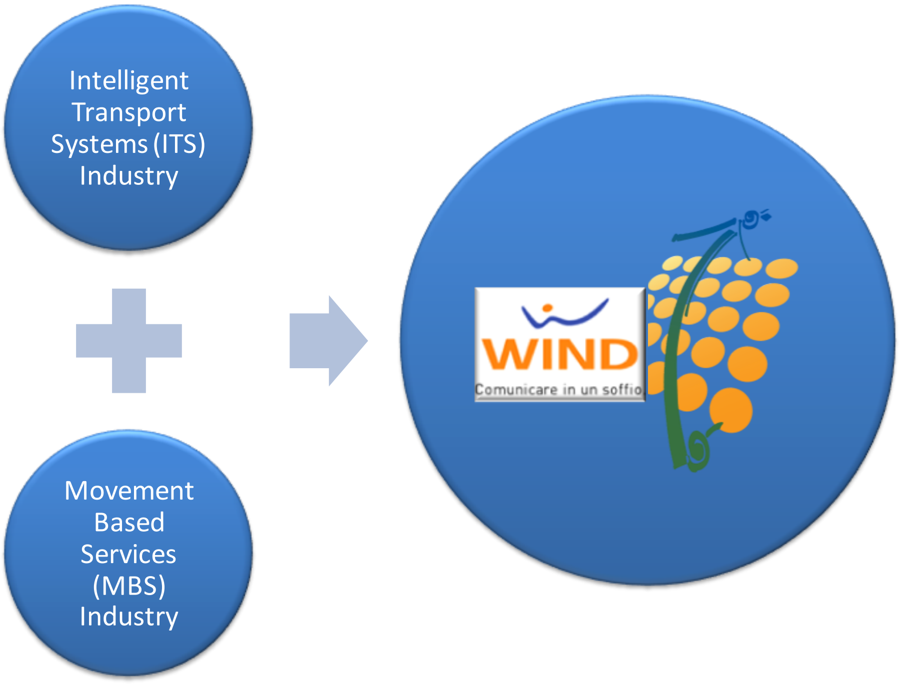


Figure 5-5 - Industry identification

Movement Based Services

This kind of service represents an innovative field where no final products are commercialized. In order to better understand what kind of tools this name refers to, a comparison with LBS services is proposed, since this last kind of service is already present in the market.

LBS stands for Location-Based Service, which can be rephrased as the solution to the question "give me some service depending on where I am now". It refers to all those services that analyze the actual position of the user through data coming from his mobile device, that can be GPS or GSM based, in order to provide real time information. One example of LBS is the service that allows logistic companies to have their fleet of trucks continuously under control through GPS devices mounted on vehicles; another example is real time information available to the final user about the current traffic situation depending on where the user is at the moment; a further example is location based advertisement that is available on a user's mobile device when he passes close to a particular place (a shop, a restaurant and so on). Other applications include services for tourism and transportation management.

These kinds of services are already present on the market and they show a great potential of investments: they have a 50 % annual market growth and they are expected to go from 144 million Euro in 2005 to 622 million Euro in 2010, according to Berg Insight (research made in 09-2006):

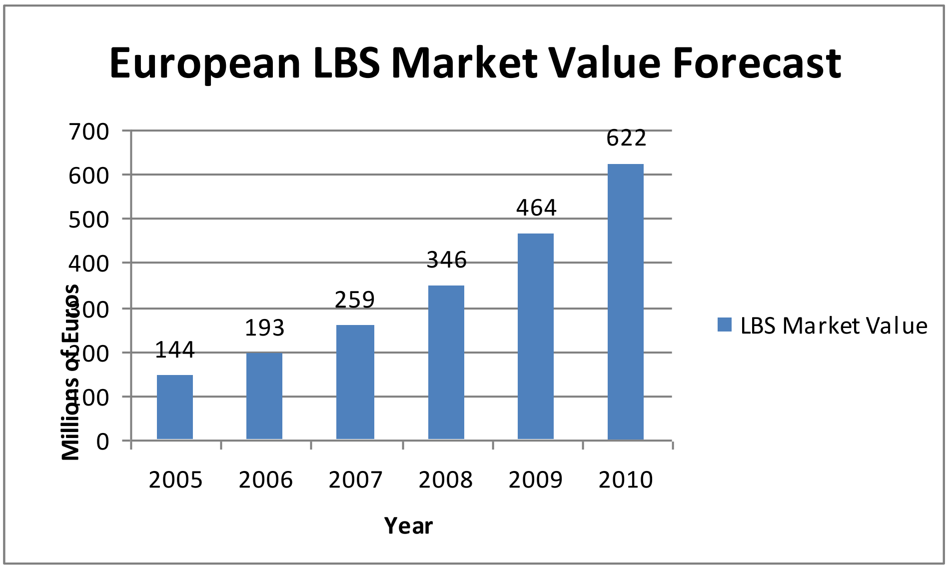


Figure 5-6 - LBS market forecast

What the project wants to offer are Movement-Based Services (MBS), where the new core information will be the cumulative movement of people, or in this case, the trajectories of the device users; the new service can then be rephrased as “give me some service depending on where I and other people have been in the past”. Therefore, starting from a historical analysis of past data, MBS services want to provide a forecast of future behavior of people. That means that MBS do not refer to real time information or with on line applications: information is analyzed not in an off line mode and it is based on the data referring to the past, not to the present.

These kinds of services are not commercialized jet, but are only at an experimental phase. A research in this field has discovered the existence of two experimental projects that are dealing with MBS services: an explanation of the areas of interest of these projects is proposed below.

- Connected Car project, by TomTom and Vodafone Netherlands: TomTom and Vodafone Netherlands are planning to jointly develop and introduce a solution that will see TomTom's technology convert anonymous, raw GSM data supplied by Vodafone Netherlands' network into accurate, real-time, information on the speed and direction of cars traveling throughout the road network of the Netherlands. The technology has been already tested and proven on a regional scale. The information will initially be available to TomTom subscribers exclusively. The Travel Time Information will be distributed using Vodafone's data network in the Netherlands. TomTom travel time information service will cover all major roads of the Netherlands as opposed to just the motorways. It will also contain door to door travel time information rather than just the length of a traffic jam, also allowing for comparison of various alternatives for total travel time and can automatically calculate the best alternative. The solution will be made available for consumers but also for road authorities and businesses, who may use it for dynamic traffic control measures and improved fleet management.
- LocHNESs by Telecom Italia: LocHNESs project, developed by Telecom Italia, uses in an anonymous way the localization if mobile devices in a certain zone, allowing to estimate traffic flow. Information about movement of devices is used to create real time maps of traffic with the analysis of the average speed on the streets and to make statistical analysis of traffic flow. This information can be used both by traffic managers and final user, who can take advantage of the calculation of travel time on a certain route, individuation of the best pattern and so on. In LocHNESs's view, it will be possible to send to users, for example by SMS, information about traffic situation periodically or on demand. User can also receive the best route, the average travel time from one point to another, taking into consideration the status of traffic in a particular moment. The system doesn't require any impact on devices or on the landscape. The localization of devices is obtained by an algorithm

owned by Telecom Italia and it is made without the identification of the telephonic number or of the user identity.

Intelligent Transportation System

In order to manage traffic and public transportation, a specific type of software is currently used, that is Intelligent Transport Systems. In the period from 1992 to around 1995 the ITS sector was known as Intelligent Vehicle Highway Systems (IVHS). At the time it was recognized that all forms of transport could benefit from the application of information and communications technologies (ICT). However the term ICT had not yet been described in popular vernacular. The global leaders in ITS at the time then determined that there needed to be a term to describe the application of ICT to transport and coined the term Intelligent Transportation Systems.

ITS is a very specialized field in which high technical competences and great investments are required. The increasing interest in ITS comes from the problems caused by traffic congestion worldwide and synergies with new information technologies; some of the possible fields of interests are simulation, real-time control and communications networks. It can also play a role in the rapid mass evacuation of people in urban centers after mass casualty events or as a result of a natural disaster or threat.

Traffic congestion has been increasing world-wide as a result of increased motorization, urbanization, population growth and changes in population density. Congestion reduces efficiency of transportation infrastructure and increases travel time, air pollution and fuel consumption.

Intelligent transportation systems vary in technologies applied, from basic management systems such as car navigation, traffic signal control systems, container management systems, variable message signs or speed cameras to monitoring applications such as traffic cameras systems, and then to more advanced applications which integrate live data and feedback from a number of other sources. Additionally, predictive techniques are being developed, to allow advanced modeling and comparison with historical baseline data.

The market of the software used in the ITS field is not so dense, since a little set of transportation management tool is currently available. The most used software are exposed below, with a brief explanation of their functionalities and features [34][35][36][37][38][39].

- ArcView 9.2: it provides geographic data visualization, mapping, management, and analysis capabilities along with the ability to create and edit data. One of the key features is the possibility of creating maps of high quality using simple wizards and an extensive suite of map elements; it is also possible to create interactive maps from file, database, and online sources that allows the user to access a wide variety of digital data, to create street-level maps that incorporate GPS locations and to view satellite images. Furthermore this tool allows to analyze spatial data and derive answers from data of a location-dependent nature, along with the deployment of GIS data in order to share and deliver interactive maps based on dynamic content. It can communicate more efficiently with the ability to graphically mark up maps and utilize ArcWeb Services in ArcReader including route, nearby place, and address finding. For all these reasons, this software is currently the most used in the urban planning and traffic management fields. The cost of this tool is 1500$ for a single license.
- CUBE: it is a family of software products that form a complete travel forecasting system. CUBE is comprised of CUBE base and add-on libraries of planning functions. This structure allows the professional planner to add functions as required without the need to learn a new interface and without the need to create multiple planning databases. CUBE is a modular system and pricing of the modules ranges from $3,500 to $12,500. Multi-seat and research discounts are available. An user of CUBE would acquire CUBE base for passenger forecasting, commodity flow forecasting, micro-simulation, matrix estimation or land use. Instead Cube Voyager combines the latest technologies for the forecasting of personal travel. CUBE Base and voyager cost $6,000 and $6,500. CUBE allows importing dBase files, Microsoft excel, comma-separated values (CSV), ASCII text files,

shape files and all major graphics formats as GIF, TIFF etc. CUBE can export to dBase, Microsoft Excel, CSV, ASCII and shape files. There are no limitations on size for the number of zones, links, nodes or transit lines. CUBE can perform trip generation using either regression, cross-classification or trip rates. Activity, discrete choice records based processing can also be implemented in CUBE. CUBE can perform traffic assignment and can use transit system based, route based, timetable based or multi-path assignment procedures. Further, it allows discrete multi-routing during the transit trip assignment. CUBE permits direct usage of ArcGIS. It also has the ability to export to other GIS packages and can be used to model and simulate pedestrian and bicycle travel.

- Emme 2: it is a travel demand forecasting software with network editing tools, visualization and analysis capabilities, GIS integration capabilities, and an extensive and extensible library of maps and charts. Emme 2 is built to handle the transport systems and addresses urban planners taking into account technological, social and economic issues. Emme 2 offers algorithms of both equilibrium assignment and integrated multimodal assignment for transport modelling. It is able to display imagery or shape-files in the background to locate node coordinates and provide visual context and It also enables comprehensive scenario comparisons to illustrate network differences at a glance. It is possible to export Emme network to shape-files for use in GIS software or Google Earth and to move data to and from spreadsheets, databases, and statistical software with CSV files or custom delimited text file formats. It is only capable of exporting to ASCII files. Emme 2 can implement trip-generation using either regression, cross-classification or trip rates. Any trip distribution model, calibration to observed trip length distribution and matrix adjustment using observed counts can be implemented using matrix balancing procedures. The listed price of a single-user license is $ 9,000 and varies with computing platform and application size. Multiple license discounts and academic discounts are also available.

- Google SketchUp Pro 5: it is an user-friendly 3D software that combines a number of easy-to-use and effective tools with an intelligent drawing system that simplifies the process of 3D design. The functionalities of this software include allowing the modelling of organic forms and the simulation of camera movements. It also allows to import models from other applications and export SketchUp models in 2D and 3D applications and also in GoogleEarth. This tool addresses in particular urban planners and all the agencies involved in streets and building designing and construction. It can not be used for running simulations on traffic flow and to analyse the movement of people on the streets. There are two versions of this software available: the basic one which is free and the professional one which costs 495$.
- Metroquest this software is a sophisticated integrated modeling system that allows users to instantly see the connections between choices and consequences as they explore alternative future scenarios. It can be a support tool for decisions about priority issues such as where people will live and work, transportation and environmental policies; it allows to visualize consequences with maps and graphs that can show the 40-year consequences of policy choices; it is able to create scenarios and watch priority indicators change, highlighting tradeoffs; it views dozens of detailed results about such varied issues as land use and infrastructure spending to transportation and air quality. It permits also to obtain input from stakeholders to establish planning priorities, to visualize alternatives and explore options using a wide range of indicators, and to evaluate future scenarios based on the community's values and priorities. It addresses fast growing cities and regions, planners promoting stakeholder engagement, decision makers embracing sustainability, municipalities focusing on growth management, and regions conducting long term planning. MetroQuest is customized for each region, allowing discussions to be focused on relevant local issues:

so it is impossible to obtain a standard price.

- Streetscape Pro: this tool is specialized in creating quick visualizations of street designs. It allows the user to keep track of numbers of on-street parking spaces, to measure height/width ratio of street design, to analyze and understand street design during the design process, and to create street designs which are compatible with the CityCAD planning software. Also this software addresses only urban designers and planners and it is available freely on the web site.
- TransCAD: it is a GIS designed specifically to store, display, manage and analyze transportation data. TransCAD combines GIS and transportation modeling capabilities in a single integrated platform. TransCAD can be used for all modes of transportation, at any scale or level of detail. It is popular in that the look and operation of the software is very GIS menu oriented. There are no limits on the number of zones, number of links, nodes or transit lines in TransCAD. This software can perform trip generation using either regression, cross-classification or trip rates. It can perform traffic assignment and is also capable of using stochastic user equilibrium, general cost assignment, multi-user and multi-class assignment, toll road assignment, assignment with intersection turn delays and signal optimization. TransCAD can use transit system based, route based, timetable based or multi-path assignment procedures and stochastic equilibrium assignment with route capacities. This software can have separate and fully integrated networks for bicycles and pedestrians. TransCAD is compatible with virtually all land use models and can be linked to them through GIS files. It can host the input and output data from land use models, display and color code parcel and land use data directly and can transform data between disparate zone systems and networks. TransCAD Standard License costs $9,995. Multiple license discounts, academic discount and discounted prices in many states and regions are also available.

- Vissim: it is a microscopic simulation program for multi-modal traffic flow modeling. It is able to simulate urban and highway traffic, including pedestrians, cyclists and motorized vehicles. Vissim addresses transportation professionals who want to simulate different traffic scenarios before starting implementation. It thus allows to find a solution which takes traffic and transportation quality, safety and cost into consideration. Vissim, though, is not only used by transportation professionals, but also by decision makers and local authorities, regardless of whether a new road is going to be constructed or a new tram line is being planned. Through this software it is possible to compare intersections with regard to design alternatives; to simulate and visualize bus/underground connections; to analyze bus and rail acceleration measures based on individual prioritization parameters, design and simulate the placement of toll roads checkpoints; to simulate dedicated lane use and shared traffic areas including all relevant road users; to carry out analyses considering alternative route control, traffic flow control, special lanes and section control systems; to simulate motorway driver behaviour and the impact on traffic flows; to create a feasibility analysis of large networks (for example motorways) with alternative route choice using dynamic assignment; and to model parking spaces with parking guidance systems, access control and dynamic level and space counting displays. Vissim base version costs $ 495, professional $ 795, and professional with customer care $ 3000.
- Visum: it is a comprehensive, flexible software system for transportation planning, travel demand modeling and network data management. Designed for multimodal analysis, Visum integrates all relevant modes of transportation into one consistent network model. It provides assignment procedures and models the peak hour which provides direct or nearly direct input for intersection capacity analysis. Similar to TransCAD, it has expansive data import and export abilities built into the software. There are no limitations in size for the number of zones, links, nodes or transit lines. Visum can

generate trips using regression, cross-classification, trip rate daily activity schedules and time of day generation methods. Visum can perform traffic assignment with multi-class (single occupant vehicle, HOV and truck) assignment models. Unique to Visum is the bi-criterion user equilibrium for road pricing analysis. This software can use transit system based, route based, timetable based or multi-path assignment procedures and stochastic equilibrium assignment with route capacities. It also allows a user definable path choice model and direct assignment of transit itineraries from survey records or electronic ticketing. Visum maintains a geographically accurate street network, including the exact shape and length of links. The boundaries of zones and higher-level area objects are maintained as part of the data set. As in GIS software, all network objects can have as many user-defined data variables as wished. It also allows additional non-network geography layers like land-use/zoning rivers, etc. Visum provides import and export ability of networks and other information to ESRI's personal geo database where the scenarios can be conflated and spatially analyzed. Walk and bike trips are included in the trip chain model as mode choice alternatives. Visum Base Version for modeling up to 400 Zones costs $5,000 and Visum X-Large Version for modeling up to 5000 zones costs $20,000. Multiple license discounts and discounts are also available.

In order to have a clearer idea of the functionalities implemented by each of the software shown above, a table is proposed with the features covered by each tool:

Software	ArcView	CUBE	Emme 2	Google SketchUp	Metroquest	Streetscape Pro	TransCAD	UrbanSim	VISSIM	VISUM
Simulation of traffic flow	no	yes	yes	no	yes	no	yes	yes	yes	yes
Multimodal simulation	no	yes	no	no	yes	no	yes	no	yes	yes
Map visualization and analysis	yes	yes	yes	no	yes	no	yes	no	yes	yes
Association of geographic data with meaningful information	yes	yes	yes	no	no	no	yes	no	yes	yes
Creation of maps	yes	yes	yes	no	no	no	yes	no	yes	yes
Design of streets	no	no	no	yes	no	yes	no	no	no	no
Forecast of future scenarios	no	yes	yes	no	yes	no	yes	no	yes	yes
Data querying	yes	no	no	no	no	no	no	no	no	no

Table 5-2 - Comparison between software

5.2 Gather information from the client

In order to gather the actual needs of the possible final users, direct contacts with a number of urban traffic managers have been established. With some of them, a face to face interview was performed, while with others a written questionnaire was carried out.

The basic questionnaire followed in order to collect users' needs is shown below:

General questions:

- What reasons led you to participate in the GeoPKDD project?
- Which is the most critical transportation mean to manage?
- What are the channels currently used in order to gather the needs of the final user?
- Which are the main investment fields you are currently dealing with in the management of each of the public means? How much financial efforts are you experiencing in these fields?

Traffic management:

- What kind of problems that you daily face in traffic management do you expect to solve with GeoPKDD implementation?
- What do you currently use to keep the traffic situation under control?
- How much is the economical effort in implementing the current tools used for traffic management? Which are their benefits and criticalities?
- Which are the sources of data used to realize traffic management (cameras, surveys to clients etc.)? Do you think they're enough representative?
- How do the meteorological conditions affects the traffic flow? How do you currently manage it?

Route planning:

- Which are the main problems you currently find in bus network management? What about metro?
- What kind of information are currently lacking in order to have a more effective route planning?

Technical questions:

- What are the features that made you choose GeoPKDD application?
- What was your first impression about the technology used in GeoPKDD such as MBS (Movement Based Services)? What do you think are the advantages and disadvantages of this technology?
- Can you provide a list of the software you currently use to plan public transportation routes? Are they web based application or desktop based application?
- What are the pluses and minuses of these software products? Do they satisfy your needs?
- What other devices except computers do you use for urban planning?
- What features do you think the prototype definitely has to have?
- What are the end users of GeoPKDD application? What are their expertise in the field of urban planning and routes planning?
- What do you expect to have from GeoPKDD application prototype?

Specific needs questions:

- Between the needs that have already emerged, which ones do you consider the more important ones?
- Would you choose a tool instead of another only because it has got this feature (referring to one in specific)?
- Would you be minded to additionally pay for a tool containing this

feature? How much is its value added?

- Would you ever implement a tool not containing this feature?

Latent needs questions:

- How much interest does your mobility agency have towards the implementation of info-mobility services for final users?
- Do you have at disposal graphical instruments to visualize traffic flow?
- Are you able to run traffic simulations? If yes, how do you do it? What kind of situations are you able to simulate?
- On which bases do you currently decide how many means to allocate to each route?
- Do you already have some tools that calculate the average travel time from one point to another? If yes, in which way and with how much accuracy? If not, do you think it could help supporting decisions about urban planning?
- Which kind of instruments do you have at disposal to foresee and avoid situations of over-crowded streets?
- How do you manage the happening of extraordinary events, such as concerts or football matches? In which way do you foresee the flow? How accurate is the forecast currently?
- In which way is currently managed the selling of advertising spaces on bus and metro network? Do you think it could be made in a more effective way?

The contacts for this gathering phase are:

- AMA, Mobility Agency of Milan;
- Comune di Vercelli, Dipartimento di Urbanistica;
- Mobility Agency of Ferrara.

In the following part, the results of the several interviews are proposed.

They are still presented in a raw way, since the managing and the analysing of them will be made in the next step of the process.

5.2.1 AMA, Mobility Agency of Milan

The interface which provided the following information is Fabio Pressi, Marketing Responsible for Infoblu - Gruppo Autostrade, who worked for several years inside AMA and is now an external consultant on info-mobility issues.

During the meeting, Fabio exposed which could be from his competent point of view the main needs of the mobility agency of Milan. First of all, he spoke about origin destination matrixes that are used from mobility agencies in order to estimate the average flow from one point of interest to another. Currently these matrixes are calculated every five or ten years with very high costs and low accuracy: he estimated that the average cost of this information is 50000 euros per year. This implies the need to have a more accurate and less expensive tool that calculates this matrix more frequently.

Then he mentioned the need to estimate the average travel time from one zone to another. Cities positioned sensors on some of the key roads in order to measure the average time that a vehicle takes to go from one point to another; the problem is that it is very expensive to create and it is not possible to understand the actual path followed by the user to go from the starting sensor to the ending one. In this context, Infoblu is able to calculate this measure through data coming from the 6 million of Telepass that send data every ten kilometres, but it is able to make it only on Italian highways and not on city streets.

Another important issue for a mobility agency and also for Infoblu is the estimation of percentage of turns on the main roads. This need derives from the fact that, even if it is possible to know the actual load of vehicles on a road, it is still impossible to foresee how many cars will take one exit street instead of another. Nowadays even with Telepass on highways this information is still unknown, but it could be very useful to have the situation of traffic under control.

At least, a crucial information that mobility agencies still lack in order to better manage the traffic flow and the urban planning is the actual impedance on the streets. Impedance is the measure of the variability of traffic load on a specific street during the day, depending on how the road is free or busy of vehicles. This value can be estimated both on the average speed on the street with respect to the maximum speed allowed or on the number of vehicles present on the road. This index can vary from 0 to 1, where 0 means that the road is blocked and 1 that the street is empty. There are already in the market simulators that estimate the variation of traffic flow on the streets, but still there is no accurate information about the actual value of impedance and of its variability depending on the period of the day and on the presence of extraordinary events. The availability of this information could help in simulating the traffic flow in a more accurate way and to better calculate the best road to undertake in order to avoid situations of traffic congestions. Furthermore it could be possible to estimate the flow of movements in presence of extraordinary events and depending on the period of the day or of the week, simulating the happening of future congestions and evaluating possible solutions to address the problem.

After the exposition of the possible users' needs and requirements, the discussion was turned into a more concrete and practical issue, that is which functionalities have to be implemented in the final application addressed to mobility agencies. Fabio pointed out two main features: origin and destination matrix from one hand, and average travel time from the other.

5.2.2 Comune di Vercelli, Dipartimento di Urbanistica

Concerning the mobility of Comune di Vercelli, a written questionnaire was submitted to the contact Arch. Franco Zanello, responsible of Dipartimento Urbanistica.

In order to gather citizens' needs in mobility issues, Comune di Vercelli is currently using updated databases created for public bikes service and for parking charging through calling cards. Other informative channels are questionnaire or face to face interviews, together with direct measurements on streets. Franco said that Vercelli is trying to invest in order to keep the situation of traffic under control through planning tools, traffic plans and so on, but with very few results. They are also trying to keep trace of the average travel time between different zones of the city, but the only instruments at disposal are computer, pencil, watch and the results of the questionnaires mentioned above: they feel the lack of effective software tools able to provide more accurate information. Some desktop-based software products are available, but they are not appropriate for the actual needs and they are not used because of lack of time; in addition, these tools are not able to run simulations on traffic flow, which instead would be really useful.

Furthermore, Vercelli feels the need to have at disposal more dynamic data in order to better follow the evolution of urban mobility: information, up to now, is not collected frequently enough and it doesn't mirror the actual situation.

Franco stated that the main problem Vercelli has to face in mobility issues is the habit of people of using cars and not public means, together with the absence of quality corridors for buses and the insufficient frequency of public means rides.

Another discussed issue was the way in which Comune di Vercelli faces the happening of extraordinary events, such as football games or concerts. Franco said that these situations are currently managed only when needed through the allocation of extraordinary parking areas linked to the city centre with public means. Daily peaks of traffic flow, especially in rush hours, are not actually managed, because they are considered unavoidable.

5.2.3 AMI, Mobility Agency of Ferrara

The direct contact which was used as interface in order to gather a mobility agency's needs was Ing. Alberto Croce, mobility manager of Ferrara.

Alberto said that one of the most important investment fields in which Mobility of Ferrara is involved is the creation and calibration of traffic simulation models. These simulations are useful to help to better plan the creation of new infrastructures and also have an idea of how the flow responds when extraordinary events happens (accidents, closure of streets and so on).

These models currently run with three types of input data:

- ISTAT census data that are available every 10 years and that provide only information about systematic movements (home-work, home-school and so on). These kinds of movements, according to Alberto, represent only 50% of the totality of urban movements;
- surveys made by Comune di Ferrara directly to a heterogeneous sample of citizens. These surveys are carried out every 3 years and involve about 1300 people. The contact is usually established by phone if the questionnaire is short, or by a face to face meeting in the case of a higher number of questions;
- automatic sensors mounted on some streets that count the number of passing vehicles and keep the traffic situation under control.

This information is used to create an important tool for traffic management, that is the origin and destination matrix. This matrix is used in order to have a clear picture of the situation of flow of people between different areas inside the city or between different cities at a regional level. Origin destination matrix provides an important part for the redaction of the Urban Mobility Plan, which is one of the main tasks carried out by a mobility agency.

In order to run simulations of traffic flow, two main software products are currently used by the mobility agency of Ferrara: VISUM, at a macro level, and VISSIM, at a micro level. These products are developed by a German university and are currently used by about 200 public administrations in Italy. They are used by traffic engineers in order to optimize the street network and to make analysis of capabilities. Through them it is possible to estimate for each street the ratio between flow and physical capacity of the road, which is called calibration, together with the total travel time of all the users in the network depending on possible changes simulated in the model.

Alberto expressed his hopes to have the possibility, through the GeoPKDD project, to automate the gathering of data for origin and destination matrixes to feed into the simulation models. Furthermore, his great interest was in the possibility of discovering alternative routes for a particular congested road and the repartition of flows on all these alternatives.

5.3 Identify and organize needs

Raw information gathered from clients was managed in order to extract an organized list of needs. The results of the questionnaires were analyzed in depth in order to find out not only the explicit needs but also the latent and implicit ones. Therefore, not only the questions to the answers were taken into account, but also a set of not-verbal elements like the tone of the voice, expressions of the face, implicit references to other issues and so on. This analysis was conducted in all the questionnaires made and then the several results were put together in an unique table. The duplicated needs were eliminated and the remaining needs were put in a formal structure. Particular attention was paid to expressing the several needs all at the same level of detail in order to allow comparisons between them.

Furthermore, with reference to the Kano matrix, all extracted needs were classified in one of these three categories:

- “A” type: excitements, that are requirements that the client doesn’t expect to find;
- “B” type: linear, that are requirements towards which the client expect an average performance;
- “C” type: basic, that are requirements whose lack can not be accepted by clients;

Taking into account the Kano categories and the information coming from the questionnaires, all needs were finally ranked basing on the relative importance. An evaluation from 1 to 9 was assigned to each need and it was explicated which client expressed the need. The results of this analysis are shown in the figures below:

Users' needs list	Priority	AMA	AMI	Vercelli	Kano
To have at disposal data more frequently	9	x	x	x	B
To have a more effective tool to create origin destination matrixes	9	x	x		B
To understand the actual path followed by users, not only the starting and the ending points	9	x			A
To estimate the percentuage of turns on the main roads	5	x			A
To estimate the actual value of impedance of the streets	7	x	x		B
To calculate variability of impedance depending on the period of the day and on the presence of extraordinary events	7	x	x		A
To estimate traffic flow depending on the period of the day or of the week	9	x	x	x	B
To simulate the happening of future congestions and evaluating possible solutions to address the problem	7	x			B
To have at disposal more representative data of actual movements of people	9	x	x	x	B
To estimate the actual value of how the flow responds when extraordinary events happens	9	x	x	x	B
To have at disposal a post service customer care for the software	3	x	x	x	C
To analyse and manage traffic situation at regional level	5		x		B
To automate the gathering of data for origin and destination matrixes	9	x	x		B
To have a huge amount of data to feed simulators	7		x		B
To discover alternative routes for a particular congested road	7	x	x		A

Figure 5-7 - Users' needs part 1

Users' needs list	Priority	AMA	AMI	Vercelli	Kano
To visualize maps of the urban morphology	9	×	×	×	C
To have more information about non-systematic movements	9	×	×	×	B
To reduce efforts (costs, time) to gather information from the client	9	×	×	×	B
To estimate calibration of each street	3		×		B
To calculate the total travel time of all the users in the network depending on possible changes	5		×		A
To obtain dynamic information on the actual flow of private mobility in the city, in order to take more well-timed decisions	7	×	×	×	B
To increase the presence of bicycle paths and pedestrian areas	1	×	×	×	C
To optimize the individuation of extraordinary parking zones in the presence of special events	3			×	A
To obtain a better allocation of parking zones and to improve their impacts on traffic flow	3			×	B
To obtain a better planning of traffic viability	9	×	×	×	B
To obtain an optimal plan and allocation of bus corridors that can reduce the average travel time of busses	3			×	A
To have at disposal an user friendly graphic interface	7			×	C
To optimize the allocation of resources on each route with the respect of users' needs	3			×	B
To identify the real movement needs of the citizens to make the mobility needs analysis process less expensive and more accurate	9	×	×	×	B
To estimate the use of public means by citizens in a more accurate way	5			×	B
To run simulations in a virtual environment about urban traffic, especially in presence of particular events	9	×	×	×	C
To obtain a more accurate calculation of the average travel time from one point to another, according to different temporal intervals and situations	9	×	×	×	B
To reduce the economical efforts in the management of street signs	1			×	A

Figure 5-8 - Users' needs part 2

5.4 Define functionalities

The initial technical deepening inside the GeoPKDD project, together with a market research within the traffic management market, led to the possibility of defining the several functionalities that can be implemented in a final service addressed to a mobility agency.

This is the area involving people mobility by private means. Since many people spend a great part of their daytime blocked in the big cities' traffic, an improvement in this area can really affect the citizens' quality of life.

Some of the most important impact areas that the final service can have in traffic management are:

- forecasts about traffic on the streets in presence of extraordinary events (football matches, concerts, etc);
- simulations of traffic trends in presence of modifications in the urban morphology;
- calculation of the average travel time from one point of interest to another;
- identification of the main arterial streets and classification of the streets based on their over-crowding risk;
- forecasts about construction site consequences in traffic;
- application and optimization of smart street lights.

In this field, it also has to be considered the effect of the final service in parking management. Examples of obtainable benefits are:

- optimal distribution of parking zones, specifically in presence of extraordinary events such as football matches or concerts;
- different parking pricing based on car confluence.

The last element is very simple to implement, since it requires changes only in the payment system; the first benefit, instead, is more difficult to put in practice since it involves important alterations in urban morphology.

A detailed list of possible technical functionalities was built, together with the evaluation of the technical feasibility of each of them. As a matter of fact, not all functionalities can be implemented with the same difficulty: for example, some of them can not be easily implemented with data provided by Wind, since GSM positioning method can not be as much accurate as GPS method. All these technical features were taken into account and were summarized in a final evaluation that was assigned to each functionality previously brought out. The feasibility values are from 1 to 9 and the results of this analysis are shown in the tables below:

Functionalities		Feasibility
Quality Bus corridors	Simulations of traffic flow in presence of certain events like the planning of new quality bus corridors	3
Uncovered patterns	Analysis of users' behaviour to discover frequent patterns not covered by public means	3
Public means optimization	Optimal allocation of the public means on the routes together with GPS tracking of public vehicles	3
Extraordinary events simulation	Simulation forecasts about the confluence on the streets in presence of extraordinary events	7
Urban planning simulation	Simulations of the traffic trend in presence of modifications in the urban morphology	7
Travel time	Calculation of the average travel time from one region of interest to another	9
Streets classification	Identification of the main arterial streets and classification of the streets based on their over-crowding risk	5
Traffic lights optimization	Optimization of smart street lights system	1
Environmental islands	Simulations of traffic flow for optimal distribution of pedestrian and cyclist zones	3

Figure 5-9 - Functionalities part 1

Functionalities		Feasibility
Parking optimization	Indication of the optimal distribution of parking zones, specifically in presence of extraordinary events, and definition of a charging model	3
Origin destination matrixes	Automatic gathering of data for the construction of the O-D matrix at a local and regional level	9
Trajectory reconstruction	Identification of the actual route taken by users, not only starting and ending points	1
Optimal route planning	Calculation of the optimal route depending on time period through historical data	7
Indicators calculation	Calculation of streets impedence, percentuage of turns, streets calibration and other indicators	5
Mobile devices information	Generation of forecast information of future traffic and public means situations available to the public through mobile devices	9
Construction sites	Simulations about construction site consequences in traffic	3

Figure 5-10 - Functionalities part 2

Merging together needs and functionalities, a House of Quality was finally created in order to understand which could be the best functionalities to implement in the final service. This kind of analysis permits to take into account how much each functionality satisfies the several extracted needs and, therefore, to evaluate the level of satisfaction that each functionality brings to the clients.

It is worth to underline that there are no tradeoffs in implementing the several functionalities, that means that there are no possible incompatibilities between functionalities. Therefore, in the “Roof” of the House there are only positive correlation, that corresponds to the situation in which the implementation of one functionality reduces the required efforts to implement the other one.

The intersections between needs and functionalities were evaluated with a scale from 1 to 9, that expresses how much each functionality satisfies the need if implemented.

Concerning planning matrix, whose main purpose is to compare how well the team met the customer requirements compared to its competitors, a benchmarking analysis was undertaken in order to value how much several available software tools respond to each users’ need. Through this calculation, it was estimated the “utility” related to each software with respect to the users needs’ point of view. It is important to highlight that the needs involved in this “right” side of the House are the only one related to the software side of the final service and that received a priority vote major or equal to 7.

The final ranking between functionalities was calculated with a mathematic model proposed by Lyman [20]. Through this method, each cross between needs and functionalities is weighted with the relative importance of the need and normalized with respect to the sum of the elements on each row. Then, for each column the technical feasibility was taken into account through the multiplication of the final result with the functionality’s feasibility level turned in the scale from 0,1 to 0,9.

The final result of this analysis is shown in the following figures:

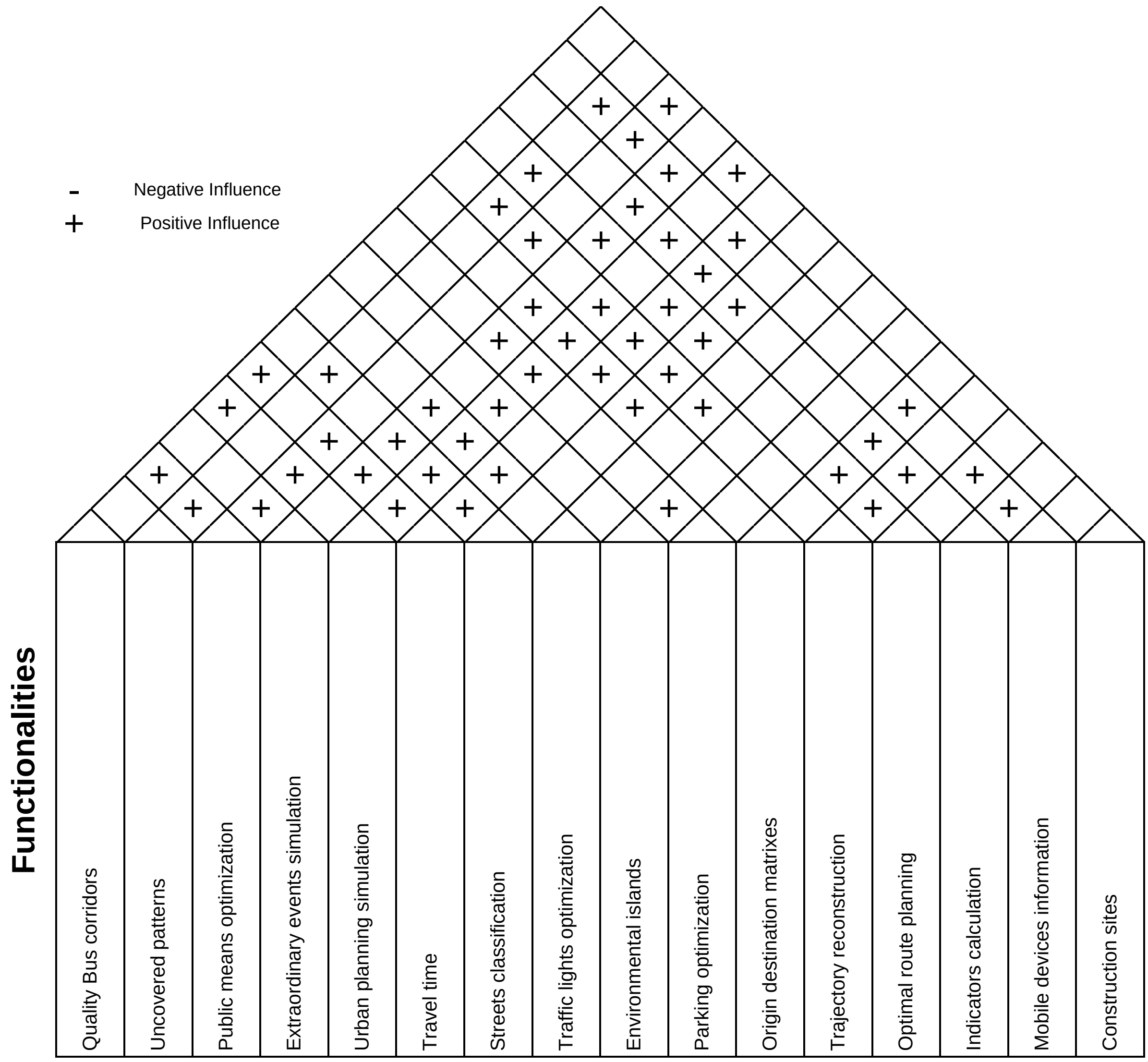


Figure 5-11 - Roof of HoQ

Users' needs list	Priority	ArcView	CUBE	Emme 2	Sketch Up	Streetscape	TransCAD	Urban Sim	Vissim	Visum	GeoPKDD
Price		**1500**	**12500**	**9000**	**495**	**0**	**9995**	**0**	**3000**	**20000**	
To have a more effective tool to create origin destination matrixes	9	0	0,1	0,1	0	0	0	0	0,1	0,1	0,9
To understand the actual path followed by user, not only the starting and the ending points	9	0	0,9	0,9	0	0	0,7	0	0,7	0,7	0
To estimate the actual value of impedance of the streets	7	0	0,5	0,5	0	0	0,5	0	0,5	0,5	0,3
To calculate variability of impedance depending on the period of the day and on the presence of extraordinary events	7	0	0,1	0,1	0	0	0,3	0	0,5	0,5	0,9
To estimate traffic flow depending on the period of the day or of the week	9	0	0,7	0,7	0	0	0,3	0	0,5	0,5	0,9
To simulate the happening of future congestions and evaluating possible solutions to address the problem	7	0	0,3	0,3	0	0	0	0	0,5	0	0,3
To estimate the actual value of how the flow responds when extraordinary events happen	9	0	0,7	0,7	0	0	0,3	0	0,3	0,3	0,9
To discover alternative routes for a congested road	7	0	0,5	0,5	0	0	0,1	0	0,3	0,3	0
To visualize maps of the urban morphology	9	0,9	0,3	0,3	0,9	0,1	0,9	0,3	0,7	0,9	0,3
To have more information about non-systematic movements	9	0	0,1	0,1	0	0	0,1	0	0,3	0,3	0,3
To reduce efforts (costs, time) to gather information from the client	9	0	0,1	0,1	0	0	0	0	0,1	0	0,3
To obtain dynamic information on the actual flow of private mobility in the city, in order to take more well-timed decisions	7	0	0,3	0,3	0	0	0,1	0	0,1	0,1	0,9
To obtain a better planning of traffic viability	9	0,3	0,9	0,9	0	0,1	0,9	0,3	0,9	0,9	0,3
To have at disposal an user friendly graphic interface	7	0,9	0,3	0,1	0,9	0,9	0,7	0,7	0,5	0,9	0,3
To identify the real movement needs of the citizens to make the mobility needs analysis process less expensive and more accurate	9	0	0,1	0,1	0	0	0,1	0	0,1	0,1	0,3
To run simulations in a virtual environment about urban traffic, especially in presence of particular events	9	0	0,9	0,9	0	0	0,7	0,1	0,9	0,9	0,9
To obtain a more accurate calculation of the average travel time from one point to another, according to different temporal intervals and situations	9	0	0,5	0,5	0	0	0,5	0	0,3	0,5	0,9
Utilità		**17,1**	**61,7**	**60,3**	**14,4**	**8,1**	**52,4**	**11,2**	**60,9**	**62,9**	**72,9**

Figure 5-12 - Benchmarking analysis

		Quality Bus corridors	Uncovered patterns	Public means optimization	Extraordinary events simulation	Urban planning simulation	Travel time	Streets classification	Traffic lights optimization	Environmental islands	Parking optimization	Origin destination matrixes	Trajectory reconstruction	Optimal route planning	Indicators calculation	Mobile devices information	Construction sites	
To have at disposal data more frequently	9		3		3	3	3	3				3			3			21
To have a more effective tool to create origin destination matrixes	9											9						9
To understand the actual path followed by users, not only the starting and the ending points	9		3					3					9					15
To estimate the percentuage of turns on the main roads	5												3		9			12
To estimate the actual value of impedance of the streets	7		1			3	3	9	3				3	3	9			34
To calculate variability of impedance depending on the period of the day and on the presence of extraordinary events	7		1		9		3		3					3	9			28
To estimate traffic flow depending on the period of the day or of the week	9		3	1		3	9	3	9	1		3		9			1	42
To simulate the happening of future congestions and evaluating possible solutions to address the problem	7	1	3		3		3	9						3		9		31
To have at disposal more representative data of actual movements of people	9					9	1	1		1		9					1	22
To estimate the actual value of how the flow responds when extraordinary events happens	9		1		9		3					3		3		3		22
To have at disposal a post service customer care for the software	3				3	3						3				3		12
To analyse and manage traffic situation at regional level	5						9					9						18

Figure 5-13 - HoQ Part 1

		Quality Bus corridors	Uncovered patterns	Public means optimization	Extraordinary events simulation	Urban planning simulation	Travel time	Streets classification	Traffic lights optimization	Environmental Islands	Parking optimization	Origin destination matrixes	Trajectory reconstruction	Optimal route planning	Indicators calculation	Mobile devices information	Construction sites	
To automate the gathering of data for origin and destination matrixes	9											9						9
To have a huge amount of data to feed simulators	7		3		9	9	3		3			3	3	3	3			39
To discover alternative routes for a particular congested road	7		9					3						9		3		24
To visualize maps of the urban morphology	9	3	3	3	3	3			3	3	3		3	3			3	33
To have more information about non-systematic movements	9		9		3	1		3						3				19
To reduce efforts (costs, time) to gather information from the client	9											3				3		6
To estimate calibration of each street	3	1	3			3	3	9	3				3		9			34
To calculate the total travel time of all the users in the network depending on possible changes	5					3	9			9	3	3					3	30
To obtain dynamic information on the actual flow of private mobility in the city, in order to take more well-timed decisions	7					9	1		9			9		9		3		40
To increase the presence of bicycle paths and pedestrian areas	1	1		1		3		1		9								15
To optimize the individuation of extraordinary parking zones in the presence of special events	3				3						9							12
To obtain a better allocation of parking zones and to improve their impacts on traffic flow	3	1		1		3		3		3	9							20

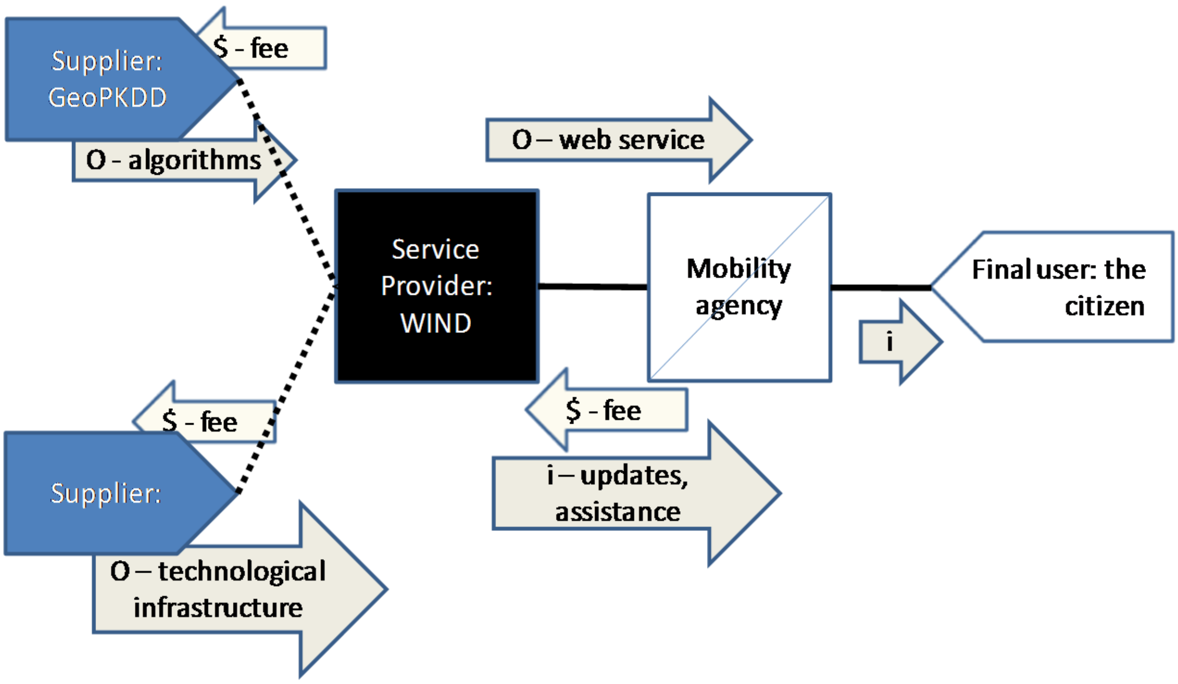
Figure 5-14 - HoQ Part 2

		Quality Bus corridors	Uncovered patterns	Public means optimization	Extraordinary events simulation	Urban planning simulation	Travel time	Streets classification	Traffic lights optimization	Environmental islands	Parking optimization	Origin destination matrixes	Trajectory reconstruction	Optimal route planning	Indicators calculation	Mobile devices information	Construction sites	
To obtain a better planning of traffic viability	9	3	3	3		9	3	3	9			3					3	39
To obtain an optimal plan and allocation of bus corridors that can reduce the average travel time of busses	3	9		3			9	3										24
To have at disposal an user friendly graphic interface	7			3	3	3					3	3	3	3		3	3	27
To optimize the allocation of resources on each route with the respect of users' needs	3	1	3	9			3											16
To identify the real movement needs of the citizens to make the mobility needs analysis process less expensive and more accurate	9					3						3						6
To estimate the use of public means by citizens in a more accurate way	5			9														9
To run simulations in a virtual environment about urban traffic, especially in presence of particular events	9	1			9			3				3				3		19
To obtain a more accurate calculation of the average travel time from one region to another, according to different temporal intervals	9					3	9		3							3		18
To reduce the economical efforts in the management of street signs	1								9									9
Technical feasibility		0,3	0,3	0,3	0,7	0,5	0,9	0,5	0,1	0,3	0,3	0,9	0,1	0,7	0,5	0,7	0,3	
Final functionalities importance		1,1	4,5	2,9	13	11	17	7,1	1	1,2	1,7	38	1	9,1	5,236	9,5	1	

Figure 5-15 - HoQ Part 3

6 Service definition for the GeoPKDD project

An accurate description of the service will now follow, where all the key elements and their relationships will be described; these include a description of the value that WIND will offer to one or several segments of customers and of the architecture of the firm and its network of partners for creating, marketing, and delivering this value and relationship capital.



6-1 The Service Business Model

The role of WIND inside the creation of the service is that of service provider, which is also its core capability. WIND, therefore, through state of the art technology and an ad-hoc technological infrastructure would be able to receive a compensation for putting at disposal of its clients the aggregated data of its user base positioning along with a web service that would allow the customer to easily access and work with this information. The customer could then indirectly pass on the benefits of the service to the final users, which are the citizens of the city, through better traffic management policies and so on.

Below is a diagram that shows how the service definition is structured:

Building Block	Description
Value Proposition	Gives an overall view of the company's bundle of services
Distribution Channel	Describes the various means of the company to get in touch with its customers
Relationship	Explains the kind of links a company establishes between itself and its different customer segments
Value configuration	Describes the arrangement of activities and resources
Core competency	Outlines the competencies necessary to execute the company's business model
Partner network	Portrays the network of cooperative agreements with other companies necessary to efficienty offer and commercialize value

Figure 6-2 The blocks of the service definition

These are the main blocks that make up the service definition, with a brief description of each. An explanation of each block will now follow, with the objective of explaining in depth the finer points of the proposed service.

6.1 Value proposition

In this part of the service definition, a description will be given of the informal notion of “the value of the service as perceived by the customer”. Two key issues must be addressed in this part:

- a description of the product itself;
- a description of the problem solved for the customer by the service.

The final service will implement the three functionalities that were identified through the House of Quality matrix, which are:

Automatic gathering of data for the construction of the O-D matrix at a local and regional level: all of the customers which were interviewed identified as a key issue obtaining more accurate data of how people move during different temporal periods, since the data that is now at disposal is either out of date, or not accurate enough. Current methods employed by mobility agencies and cities all around Italy consist mainly in the gathering of ISTAT census data that are available every 10 years and that provide only information about systematic movements (home-work, home-school and so on), surveys made directly to a heterogeneous sample of citizens, which are usually carried out every 3-5 years, and automatic sensors mounted on some streets that count the number of passing vehicles and keep the traffic situation under control. On the other hand, the kind of data that WIND can provide from its network is aggregated handover data by the hour, which records the number of active calls that switch from one cell to another adjacent to it, and also the sanitized billing card information, which includes the starting cell of the call, the ending cell, along with the total time call duration. Even though the accuracy of this data is quite raw, it can definitely provide insight regarding the origin destination problem. The city regions for which the origin destination matrix provides aggregated traffic flow can be closely approximated by the network cells and their coverage areas. Therefore, the aggregated handover data can be considered a representative sample of all the city users. The origin destination matrix, in turn, will then be used as an input for the models that are put at disposal through the state of the art technology. Up to now, the situation was analysed for urban areas, but the potential customers include also regions. In this case, the mobility needs and data are different, in that the scope is much broader and the regions of interest are much more ample. Also in this case, the kind of data that WIND can put at disposal can give a pretty accurate idea of the actual traffic flows between regions, since in this case the available data includes not only aggregated handovers, but also the location area data: the great advantage that fits perfectly with the case at hand is that, at this level, it is possible to take into account not only active calls, but also all the inactive devices that are turned on; in this way, the sample of users is considerably more accurate than before. From WIND's point of view, this kind of data is already available from the network, and its only effort would consist in the actual downloading of the data for its clients.

From the customer point of view, using more accurate information provides a much more effective and cheaper way to estimate, and therefore manage, traffic. It's also true that this kind of data is not completely accurate for a number of reasons, and in fact can only constitute a sample of the citizens. In fact, the only data that is recorded is that of active calls, which means that only someone who is calling is taken into account. What's more, only WIND users are recorded, which raises the question whether WIND user base can be considered representative in respect to the whole population. The different telecommunications companies that operate in Italy seem to have positioned themselves so that each one of them is dominant in a particular market segment, so this problem is very real, and additional studies are advised in order to clear up this doubt. Finally, handover data itself can sometimes be faulty, because a user could be on the border of two adjacent cells making a call, and multiple call handovers could take place while the user hasn't actually moved in space. All these problems require further studies in order to confirm how much they could affect the quality of the data provided to the customers.

Calculation of the average travel time from one region of interest to another: this is the other important need expressed by all the clients that were interviewed during the information gathering phase of the project. Origin destination matrices only supply the aggregated flow of users from one region of interest to the other, but the key issue that is not addressed is the amount of time needed to get from region A to region B. In order to calculate this very important measure, a number of other factors must be taken into account, and this is where the mobility agency can really make good use of the GSM billing card data provided by WIND.

In the models used to analyze traffic, the city can be mapped through a network graph, where the nodes represent the centroides of the region and the arcs stand for the roads that connect one region of interest to another; a measure is associated to each arc that is called the impedance of the road, defined as the flow of users on the road in respect to the overall capacity of the street itself.

Through the analysis of the billing cards information, which include call starting cell, call ending cell, and time of the call, it is possible to create a model that not only estimates the flow of traffic in different time intervals, but also the average travel time needed to move from one region to another, wherever it may be. Through the integration of these two types of data, it is possible to obtain a more refined estimation of how people move on an aggregated level. The results can then be seen on the GIS visualization tool approximating regions with the cells that cover that particular area. Having at disposal the timing data along with a GIS map of the area, it is possible to estimate which of the possible paths were taken by the users, and therefore assign an impedance value to each of the streets in the map.

Simulation forecasts about the confluence on the streets in presence of extraordinary events: many software tools are available on the market today that already implement some sort of traffic simulation; they are usually very expensive and not very accurate, but are nevertheless in high demand because of the importance that cities and mobility agencies can derive from them in understanding how to effectively plan an extraordinary event such as a concert or a football match. In the case of the proposed service, it could be possible to query the WIND databases in order to obtain data relating to a similar event that has happened in the past. With this particular type of data, it can be possible to estimate how the user base will behave when a similar event will take place again.

Therefore, the added value of the presented service is the possibility of working with data relating to particular events, that is non systematic ones; no other software currently available in the market is able to do this with the same accuracy as the GeoPKDD service could.

Again, it would be possible to extrapolate this added value with little economical effort from WIND’s side, since it would only be a matter of refining the query for input data that would take into account only the particular time periods in which similar events took place.

6.2 Distribution channels

The distribution channels are the means that the company can rely upon in order to get in touch with its customers. There are two main distribution channels through which WIND must act in order to guarantee the success of the proposed service, and they will now be analysed separately.

The main channel is the internet, through which the service is delivered via remote access; the alternative to delivery through the web would be to sell a retail product which would include all the different software packages developed by the GeoPKDD consortium, but this would bring about significant distribution costs and the inability of having at disposal current data from WIND's network.

The service has a web-based graphical user interface where a user can log in by inserting username and password. Once inside the program, it will be possible to access all the software tools installed on a central server that will enable the user to run a number of functionalities, such as:

- loading data in regard to format specifics,
- visualizing the data in a number of ways with an overlay of a GIS map of the region of interest,
- querying data based on the regions of interest, the specified temporal periods, and so on.

Keeping the tools on a central server has many advantages, both from the point of view of cost efficiency where the property of the software is safeguarded in the best way, and functional efficiency, since it is much simpler to maintain the server tools up and running. In addition, the software updates can be implemented very easily and there would be a single access point to WIND's network database, thereby reducing the amount of computational effort that the database could potentially have to sustain if multiple users tried to download data from retail software concurrently. The disadvantage of choosing this distribution channel is a fair investment in computational power from the server side, since it should potentially be able to manage multiple users simultaneously running software programs and querying the database. If and when the user base becomes substantial, the server could be subject to a serious strain, and therefore should be planned accordingly.

The second distribution channel is a telephone service through which it would be possible to contact clients for a number of reasons, such as quality controls, interviews, understanding which potential functionalities to add in the future, and so on. Since WIND is a telecommunications company, this should present no significant problem, other than the cost of training personnel specifically for this kind of service.

6.3 Customer relationship

In this part of the service definition, the relationship that WIND wants to establish between itself and its different customer segments should be made clear.

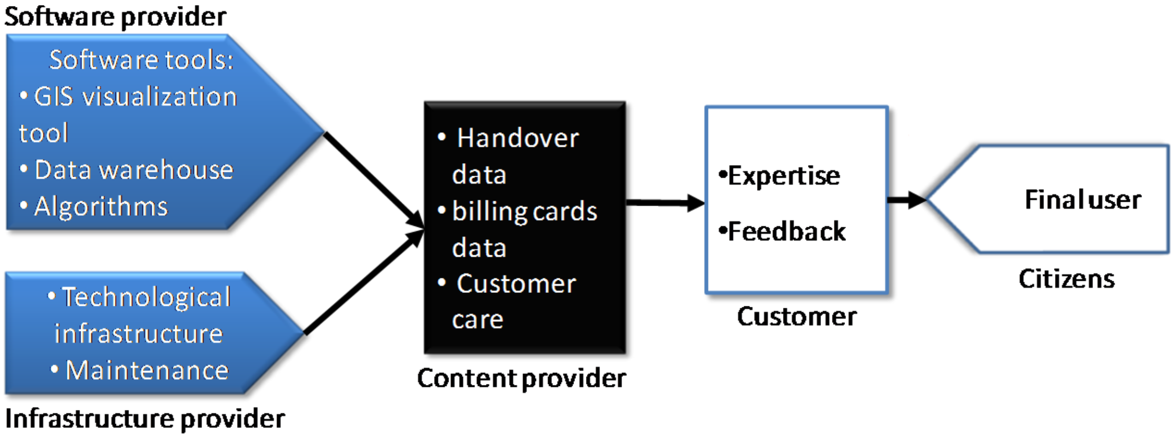
Starting from the market analysis that was already presented in this document, it is possible to plan an optimal customer relationship strategy. This service targets a potential market estimate of 17 customers; this can be a good opportunity to really focus on providing a highly customized relationship to these clients, since it is probable that they would require a lot of service advice, especially during the initial launch phase of the service. In this way, the constraint of playing in a small potential market can be turned around in an opportunity of providing a client-focused service, thereby valorizing the WIND brand in this new field.

In order to achieve this high quality of service, several actions must be taken. The most important thing is to set up a competent assistance service, that must be able to help the customer with any kind of problem that he could have, both from the technical point of view (problems with data, errors with the software tools) and from a functional point of view (help with running a particular type of analysis and so on).

On the other hand, collecting feedbacks from the customers will be the other key issue in order to guarantee the future success of the service. Therefore, it will be advisable to contact clients in order to record the possible complaints they might have, to let them know about the software updates that were made (via email), to ask which additional functionality they might like to see implemented; all of these actions can be summed up under the term of quality control.

6.4 Value configuration

This paragraph describes the arrangement of activities and resources. The figure below has the scope of highlighting the different phases that make up the service delivery process.



6-3 The Value Configuration Process

This is the scheme that was thought for this particular service, which will now be explained in its inner workings. Let's start the analysis with the two input blocks: the software provider and the infrastructure provider.

From the GeoPKDD point of view, the key elements are the three software blocks that have to be provided, which are the trajectory data warehouse, the clustering algorithms, and the GIS visualization tool; the GeoPKDD project puts at disposal this licensed state of the art technology, which will be installed on a central server that can accessed via remote; at the same time, it also must provide technical assistance for any technological problem that might come up, along with periodical software updates. In exchange for the use of these products, it will be paid a fixed fee on a certain temporal basis (month, year).

For the technological infrastructure provider, the idea would be to fix the initial contract terms with WIND concerning how to set up the server, access points, and so on. From WIND’s point of view, there are two possible alternatives that differ among themselves based on the economical effort needed to set it up: on one hand, it could either completely outsource this part of the service to an external company that would be in charge of the acquisition, set up, and maintenance of all the physical and technological properties of the network; on the other hand, it could keep it internal, and therefore allocate part of its own resources to the task of setting up and providing maintenance to the servers and so on. The choice depends on the overall cost of the two alternatives, since there is little added value in keeping this part of the service internal to WIND.

The core function for WIND once it has obtained all the necessary inputs to run the service is that of content provider, specifically providing the aggregated handover data for a particular region of interest and temporal interval as indicated by its client, along with the corresponding billing cards information. This is the real added value that WIND can provide to its customers. Through this data, it can be possible to create ad-hoc origin destination matrices that take into account the raw data coming from the cells customized to represent the different areas of interest indicated by the mobility client. This would then be one of the possible outputs that the WIND block of the diagram provides to the customer. The other function that WIND would take on is the assistance service offered to the client, along with the prompt communication of eventual software updates and implementation of new features and functionalities to the software package.

WIND can therefore integrate the different parts taken from the suppliers with its own data in order to offer the final service to its customers. The other key block in the proposed diagram is the mobility agency (or Comune or region). In exchange for the two-part fee, the mobility agency will gain access to both the software and the data, for which a series of functionalities can be implemented; the main advantage that was already mentioned is the possibility of obtaining up-to-date origin destination matrixes coming from a sample of the population (the WIND users making calls) that can then be used in conjunction with the billing cards data to calculate the average aggregated travel time between two regions.

The mobility agency is not the final customer, even though it represents it indirectly. Through access to the kind of data that WIND can put at disposal, the benefits that derive from a more efficient traffic management can then be passed on to the users through a number of initiatives, along with a better every day driving experience.

6.5 Core competency

This section has the goal of outlining the competencies necessary to execute the company's business model.

These are the three key competences without which the service cannot work:

- access to state of the art technology: the ownership of this part of the process belongs to the GeoPKDD consortium, and it will put at disposal the three technological deliverables of the project, which are the trajectory data warehouse (HERMES developed by University of Piraeus), the spatio-temporal data mining algorithms (developed by ISTI-CNR), and the trajectory visualization tool (CommonGIS developed by Institute of Fraunhofer);
- a solid technical infrastructure that has to be able to sustain multiple client analyses simultaneously: this part will be taken care of by the second key supplier to the process, which still hasn't been identified up to now. The goal of this competency is to allow correct access to the information that the client needs, and ensuring that everything works fine in order to guarantee a satisfactory service;
- access to aggregated handover data for a particular region and time period along with relative billing cards information: this is probably the most important of the three core competences, and its ownership belongs to WIND. It's only through this valuable information that the value can be created and delivered, because through the historical analysis of the data it is possible to obtain an estimate of the traffic flow that is much more accurate and up-to-date than the methods

that are currently used.

It's important to underline that, though each one of the competences is necessary, none of them could work without the others, and therefore the key issue to success must be the tight integration between the partners participation in the value creation chain.

6.6 Partner network

The scope of this part of the service definition is to describe the network of cooperative agreements with other companies necessary to efficiently offer and commercialize value.

The three key partners that will work along WIND in the creation of the service are:

- the GeoPKDD consortium;
- technology supplier and maintenance provider;
- customers: mobility agencies, Comuni, regions of Italy.

A further analysis of each one will now follow, along with their responsibilities:

- GeoPKDD consortium: this consortium is made up of 8 partners: 7 research institutes all over Europe, and 1 industrial partner, which is WIND. It is responsible of this project which is financed by the European Community, and that will end in December 2008. The relationship with WIND will consist in an exchange of the right to use the software developed by the researchers for some sort of fee, which can be paid one-time or periodically. This can be decided between the two parts in a specific contract that will have to be agreed by both before the software tools can be set up for the service;
- technology supplier and maintenance provider: this is the other supplier in the service creation process. In this case, it might be

useful for both parts to agree to an outsourcing of the technical equipment setup for the service, and this would be the role of the partner. To further define its responsibilities, it should set up the technical infrastructure needed to run the service, and this will include the servers, the connection between the service and the data warehouses, the access interfaces, and so on. On the other hand, it should also be responsible for maintaining the infrastructure, making sure that everything runs smoothly and run checks on the ability of users to access data quickly and precisely. This last issue would probably be the key one, since this would be one of the most important quality drivers on which the service could be measured, and therefore the necessary bandwidth channels must be allocated accordingly. In the same way as before with the consortium, the partnership should involve an exchange between a product (in this case intangible) that involves the correct setup, functioning, and maintenance of the network, and a fee which should be agreed upon among the two parts;

- customers: mobility agencies, Comuni, regions. The customers can also be considered as partners in the value chain of the service definition, and for this very reason they assume a very important role; with their indications and feedbacks it can be possible to deduce how well the service is being received, the future steps to take in order to fix the current problems and better the future service. In this case, it is WIND that offers a product to the customer in exchange for a price; the envisioned price is made up of two parts: one fixed part covering the right to use the software via web service, and the other periodical fee that includes both access to data for particular regions and also the maintenance service offered by the partners involved in the implementation of the service.

Conclusions

Through the several steps followed in the methodology proposed, this work provides to WIND the conception and the definition of an innovative technology push service that can take advantage of the technology developed inside the European GeoPKDD project, where WIND is the only commercial partner among the several others.

Basically, the added value that this work gives to WIND is a more market oriented view inside the technology based GeoPKDD project. As a matter of fact, before the beginning of this work, the several researchers were concentrated only on the development of the particular technological area for which they are competent: there was a lack among the partners of a more business oriented point of view and of an coherent vision about a concrete application that could be able to integrate the several specific research areas.

The reason why WIND ordered this work is to provide a business view that could justify its investments in the project and that could exploit all the opportunities that the technology at disposal offered. With this final goal in mind, the development process started, with the main peculiarity that the technology was already chosen, so a technology push development model was followed.

Starting from an in depth analysis of the state of the art technology, including spatio temporal data mining algorithms, mobile devices localization systems and trajectory visualization tools, a lot of creative ideas were brought out in order to generate new concepts for a profitable service that could be actually implemented, taking into account opportunities and constraints of the technology at disposal. These concepts were presented to the rest of the European partners during a GeoPKDD Consortium meeting that took place in Antwerp from 7^{th} to 11^{th} of October. During this meeting, all involved partners showed their satisfaction for the work done and provided some useful and competent feedback that was taken into account for the further steps of the work. Then, the several ideas were evaluated and the final concept to implement was selected. Furthermore, the new service was defined, though the identification of the functionalities that could better satisfy the users' needs, and contextualized, through a market analysis and segmentation.

What this work leaves to WIND and to the following activities of the GeoPKDD project is a marketing oriented perspective and the identification of the detailed functionalities that have to be implemented in the final application in order to satisfy customers' requirements and to create a service appealing to the market.

At the end of the whole process it was found that the service that better fits the technological constraints of the developed algorithms and of data at disposal from WIND's network is the traffic management tool. This service was defined through direct contacts with some of the addressed clients in order to identify what are the actual needs of the target market. Then, starting with the identification of these concrete needs, the functionalities to implement were chosen and, through an in depth market analysis, the business potentiality of this service was tested. These are the results that this work leaves to the further works.

It is now possible to reason on the future development and potential of this service. In the short term, the mobility agency of Milano (AMA) will start a test trial of the service in order to prove the validity of this service through the use of the aggregated handover data. The city of Milano is planning to introduce the payment of a ticket for entering with a vehicle inside the center of the city from January 1st, 2008. AMA's idea is to use the WIND data from a typical week before this date and another week after the start of this pricing policy; through the comparison of these two datasets, the representativeness of the data will be tested extensively. This will be the first real test that the application will have to pass before any further actions can be taken.

Furthermore, another important project where the application will take part is the infobrokER project, sponsored by five mobility agencies of the region Emilia Romagna in order to obtain public funding from the ELISA project for their city's info-mobility services for the final customers. Inside this proposal, the GeoPKDD application will be proposed as an innovative technology service. The key issue that is being publicized as the real added value that this service can bring is the automation of the process of calculating the origin-destination matrices, with particular emphasis at the regional level.

From a long term perspective, the service that was defined can only be a starting point to gauge the potential that could be reached; as of now, the only kind of data that can be easily utilized and that is at disposal has a rather low localization accuracy, and that is the CellID technique; in the future, more accurate technology will surely be at disposal: some telecom companies in Italy are investing heavily in this field, and it is commonly believed that GPS will become much more common in the future; this can be the starting basis to tighten an ongoing collaboration between telecommunication companies and GPS device makers, creating the A-GPS localization technique. In this way, it will be possible to obtain the accuracy of GPS together with the high diffusion that GSM currently boasts. When this technology will be available, then a whole new series of service will be possible, starting from a more accurate route planning, to the identification of the aggregated use of the single roads on the road network.

The main issue that WIND must keep into account is to move before hand in order to create partnerships with whoever will be a lucrative partner in the future. When GPS will be commonly available, WIND's GSM data will probably be outclassed, and therefore the top priority from a strategic point of view must be to work on the integration between the different types of data. In this way, WIND will not lose the market that it is trying to create with the service that has been defined inside this work. In this context, a potential partner that could be worth investigating is Octotelematics, an Italian company that builds the majority of the GPS devices present in the Italian market to date. This company also offers a number of web services for its clients that already analyze their movements with an astounding accuracy. The partnership between WIND, GeoPKDD, and Octotelematics could surely be an element to take into account for future investments.

In conclusion, it can be stated that the business oriented vision was finally provided to the GeoPKDD project, so that the technical researchers can know how to better concentrate their efforts and how to develop something that can actually meet the market's requirements.

Bibliography

[1] Bellandi G., 2006, “Il marketing engineering del prodotto”, SEU

[2] Chiesa V., 2001, “R&D strategy and organization: managing technical change in dynamic context”, Imperial College Press

[3] Abetti P.A., 1989, “Technology: a key strategic resource”, Management Review

[4] Freeman C., 1976, “Economics of industrial innovation”, Pinter Publisher

[5] Roberts E.B., 1988, “Managing invention and innovation”, Research Technology Management;

[6] Friar J. and Horwitch M., 1986, “The emergence of technology strategy: a new dimension of strategic management”, Pergamon Press;

[7] Porter M.E., 1985, “Competitive advantage”, The Free Press;

[8] Porter M.E., 1980, “Competitive strategy: techniques for analysing industries and competitors”, The Free Press;

[9] Hax A.C. and Majluf N.S., 1984, “Strategic management: an integration perspective”, Prentice Hall

[10] Hax A.C. and Majluf N.S., 1991, “The strategic concept and process: a pragmatic approach”, Prentice Hall

[11] Floyd C., 1997, “Managing technology for corporate success”, Aldershot

[12] Tidd J., 1997, “Managing innovation: integrating, technological, market and organization change”, J. Wiley

[13] Von Braun C.F., 1997, “The innovation war”, Prentice Hall PTR

[14] Bishop G, 2004, “A review of technology push product

development models and processes", ASME

[15] Bishop G., 2004, "A comprehensive model for technology push product development", Brigham Young University

[16] Garcia R. and Calantone R., 2002, "A critical look at technological innovation typology and innovativeness terminology: a literature review", The Journal of Product Innovation Management

[17] Hartmann G. and Myers M., 2000, "Technical risk, product specifications, and market risk, Managing Technical Risk", NIST GCR

[18] Miller W. and Morris L., 1999, "4th Generation R&D: managing knowledge, technology, and innovation", John Wiley & Sons

[19] Hopkinson A. J., 2007, "Evaluating the application of modularity to reduce market risk in technology push products", Brigham Young University

[20] Ulrich K. T., Eppinger S. D., Filippini R., 2007, "Progettazione e sviluppo prodotto", McGraw-Hill

[21] Terninko J., 1997, "Step-by-Step QFD customer driven product design", St. Lucie Press

[22] Paul R. N., 1987, "Improving the new product development process Making technology push work", The Journal of Business and Industrial Marketing

[23] Spivey W., Munson J., Nelson M. and Dietrich G., 1994, "Coordinating the technology transfer and transition of information technology: a phenomenological perspective", IEEE Transactions on engineering management

[24] Souder W., 1989, "Improving productivity through technology push", Technology Management

[25] Larsen J., 2001, "An engineering approach for matching technology to product applications ", Brigham Young University

[26] ISTI-CNR, 2007, "Geography, mobility, and privacy: a knowledge discovery vision", Pisa, Italy

[27] GeoPKDD Consortium (EC), 2007, "GeoPKDD: Annex I – Description of Work"

[28] www.telecomitalia.com, consulted on 1st August 2007

[29] www.esriitalia.it/lbs/index.htm, consulted on 1st August 2007, consulted on 1st August 2007

[30] www.mobility.tomtom.com, consulted on 1st August 2007

[31] www.gis.com/showcase/business.htm, consulted on 1st August 2007

[32] www.locationet.com/LBS_app/intro.shtml, consulted on 1st August 2007

[33] www.directionsmag.com, consulted on 1st August 2007

[34] cuspa.washington.edu/, consulted on 1st November 2007

[35] www.esri.com/software/arcgis/arcview.html, consulted on 1st November 2007

[36] sketchup.google.com/product_sup.html, consulted on 1st November 2007

[37] www.holisticcity.co.uk, consulted on 1st November 2007

[38] www.envisiontools.com, consulted on 1st November 2007

[39] www.tpsitalia.it, consulted on 1st November 2007

Figures and tables

Figure 1-1 - Types of Development Models 10

Figure 1-2 – Linear technology push model.. 11

Figure 1-3 - Linear market pull model 11

Figure 1-4 - Combination of technology push and market pull model 12

Figure 1-5 - Technology push six phases model 17

Figure 1-6 – Stages of technology push model related to market pull model

.... 17
Figure 1-7 - 8 steps model 18
Figure 1-8 – TAS process 19
Figure 1-9 - The dimensions of technology strategy 21
Figure 1-10 - Types of technology strategies 25
Figure 2-1 - Logic sequence of activities 28
Figure 2-2 - Number of ideas 29
Figure 2-3 - Costs during development 31
Figure 2-4 - Three semantic areas in the generation process 39
Figure 2-5 - Bubble Diagram 44
Figure 2-6 - Service definition process 45
Figure 2-7 - Ansoff matrix 46
Figure 2-8 - S.W.O.T. analysis 49
Figure 2-9 - Kano matrix 56
Figure 2-10 - House of Quality 59
Figure 3-1 - The five Work Packages 65
Figure 3-2 - Visualization of GeoPKDD process 69
Figure 4-1 - Three groups of ideas 93
Figure 4-2 - Legenda 93
Figure 4-3 - Brainstorming output part 1 94
Figure 4-4 - Brainstorming output part 2 95
Figure 4-5 - Brainstorming output part 3 96
Figure 4-6 - Brainstorming output part 4 97
Figure 4-7 - Screening process part 1 99
Figure 4-8 - Screening process part 2 100
Figure 4-9 - The completed bubble diagram 109
Figure 5-1 - Ansoff matrix 110
Figure 5-2 - Public transportation companies market analysis 111
Figure 5-3 - SWOT analysis for WIND 115

Figure 5-4 - Market segmentation 118
Figure 5-5 - Industry identification 120
Figure 5-6 - LBS market forecast 121
Figure 5-7 - Users' needs part 1 139
Figure 5-8 - Users' needs part 2 140
Figure 5-9 - Functionalities part 1 143
Figure 5-10 - Functionalities part 2 144
Figure 5-11 - Roof of HoQ 146
Figure 5-12 - Benchmarking analysis 147
Figure 5-13 - HoQ Part 1 148
Figure 5-14 - HoQ Part 2 149
Figure 5-15 - HoQ Part 3 150
6-1 The Service Business Model 151
Figure 6-2 The blocks of the service definition 152
6-3 The Value Configuration Process 158

Table 3-1 Participant list 82
Table 3-2 - LBS and MBS services 89
Table 3 - Participants to brainstorming session 92
Table 5-1- Functionalities of some mobility agencies 113
Table 5-2 - Comparison between software 130

Ringraziamenti

Doverosi e sentiti ringraziamenti alla mia famiglia. Grazie per avermi così calorosamente sostenuto in tutti questi cinque anni universitari e per aver sempre creduto in me, anche più di quanto io stesso abbia mai fatto. Non so se saprò mai ripagarli di tutti i sacrifici fatti, ma so che un modo per iniziare è renderli fieri di me e mi auguro, con tutto il cuore, di esserci finora riuscito.

Cari ringraziamenti anche al Consorzio ELIS, per avermi offerto l’opportunità di questa splendida esperienza che mi ha maturato sia professionalmente che personalmente. Ringrazio in particolar modo il mio team leader, per il sostegno e la stima che mi ha sempre dimostrato, e i miei compagni del JC08, per aver reso questa avventura così speciale. Grazie inoltre al professor Bellandi per i preziosi consigli e per la disponibilità dimostrata nel seguirmi durante la realizzazione di questo lavoro.

Ultimi, ma non ultimi, i miei amici più cari, quelli che mi hanno incoraggiato e accompagnato nei momenti più belli della mia vita universitaria. Grazie a tutti loro, per aver reso questo periodo che si chiude oggi così indimenticabile.