FROM DELPHI TO SCENARIO BY USING CLUSTER ANALYSIS: TURKISH FORESIGHT CASE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BAŞAK SAKARYA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
SCIENCE AND TECHNOLOGY POLICY STUDIES

MAY 2007

Approval of the Graduate School of Social Sciences

Prof. Dr. Sencer Ayata
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science/Arts / Doctor of Philosophy.

Assoc. Prof. Dr. Erkan Erdil
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science/Arts/Doctor of Philosophy.

Prof. Dr. Haluk Geray
Supervisor

Examining Committee Members

Prof. Dr. Haluk Geray

Assoc. Prof. Dr. Erkan Erdil

Assist. Prof. Dr. M. Teoman Pamukçu

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Başak Sakarya

Signature :

ABSTRACT

FROM DELPHI TO SCENARIO BY USING CLUSTER ANALYSIS:
TURKISH FORESIGHT CASE

Başak Sakarya
M.S., Science and Technology Policy Studies
Supervisor : Prof. Dr. Haluk Geray

May 2007, 119 pages

In this thesis, the technologies that appeared to be strategic according to the Vision 2023 Technology Foresight Project were examined in terms of how they might form up technology clusters. This thesis aims to identify technology clusters in terms of common knowledge base and to use these clusters in future scenarios as a foresight tool. In this study, Vision 2023 Delphi survey respondents' intersecting expertise levels in different fields were accepted as indicators of common knowledge base in these fields and technology clusters were formed up in this direction. In order to attain technology clusters, the appropriateness of hierarchical and nonhierarchical clustering methods and projection techniques were examined. Taking the clusters into consideration, Ward's method revealed the healthiest results for our data set.

Investigation of scenario building which had not been used in Turkey as a an effective foresight tool, forms the second step of this study. Scenario method was examined from a historical perspective and different approaches were investigated. Finally, using the technology clusters that were gained through Ward clustering, a scenario building study by scenario matrix was conducted as an example.

Keywords: Cluster Analysis, Foresight, Delphi, Scenario Building

ÖZ

KÜMELEME YÖNTEMİ KULLANARAK DELFİDEN SENARYOYA:
TÜRKİYE ÖNGÖRÜ ÇALIŞMASI

Başak Sakarya
Bilim ve Teknoloji Politikası ve Çalışmaları
Tez Yöneticisi : Prof. Dr. Haluk Geray

Mayıs 2007, 119 sayfa

Bu tezde, Vizyon 2023 Teknoloji Öngörü Projesi sonucunda ortaya çıkan stratejik teknolojiler, oluşturabilecekleri teknoloji kümeleri açısından incelenmiştir. Bu tez, ortak bilgi tabanına dayanarak teknoloji kümelerinin belirlenmesini ve bu kümelerin bir öngörü metodu olarak gelecek senaryolarında kullanılmasını amaçlamaktadır. Bu çalışmada, Vizyon 2023 Delfi anketine katılanların farklı alanlardaki birbiriyle kesişen uzmanlık dereceleri, bu alanlardaki ortak bilgi tabanının göstergesi olarak kabul edilmiş ve teknoloji kümeleri bu doğrultuda oluşturulmuştur. Kümeleri elde edebilmek için hiyerarşik ve hiyerarşik olmayan kümeleme teknikleri ile projeksiyon yöntemlerinin kullanılabilirliği incelenmiştir. Bulunan kümeler gözönüne alındığında Ward metodolojisinin en sağlıklı sonucu verdiği belirlenmiştir.

Türkiye'de henüz kullanılmayan bir öngörü yöntemi olan senaryo yazımının incelenmesi, bu çalışmanın ikinci ayağını oluşturmaktadır. Senaryo yöntemi tarihsel olarak incelenmiş, farklı yaklaşımlar üzerinde durulmuş ve senaryo matrisi oluşturulması yöntemiyle, Ward kümelemesi sonucu elde edilen teknoloji kümeleri kullanılarak örnek bir senaryo yazım çalışması gerçekleştirilmiştir.

Anahtar Kelimeler: Kümeleme Analizi, Öngörü, Delfi, Senaryo yazımı

To My Mother

And

My Grandmother

ACKNOWLEDGMENTS

This thesis is not an accomplishment of the solitary works of its author; it was developed through the support, guidance and efforts of a number of people.

In the first place, I would like to express my sincere gratitude to my supervisor Prof. Dr. Haluk Geray for not only his valuable guidance and support as an instructor, but also for his understanding and trust. I am also grateful to the examining committee members Assoc. Prof. Dr. Erkan Erdil and Assist. Prof. Dr. M. Teoman Pamukçu for their important contributions and compassion.

I owe many thanks to my friends who never left me alone throughout the study. I am indebted to Emek Akman and Derya Fındık for sharing their valuable time and knowledge with me and for their dear friendship. I would also thank Dilek Güler and Semih Yılmaz for being friends, rather than co-workers and for their patience. I wish to thank my friends Evrim Engin, İdil Gaziulusoy and Sibel Konu for supporting me even from such a long distance. Special thanks to Ahmet Haşim Eşrefoğlu for his kindness and confidence in me, he made a new city a home for me.

To Deniz Akhanlı I would like to express my deepest thanks. He endured my crazy times and shared my excitement. I could not have completed this study or change the direction of my life without his precious help, continuous encouragement and invaluable guidance. He has been the music in my mind and my heart.

My greatest debt is to my family. I am grateful to my father Osman Sakarya for being such a caring, unique father and I am grateful to my brother Gürak Sakarya for his endless support, he has been the source of joy whenever I felt depressed. This thesis is dedicated to my mother and my grandmother for the love and strength they inherited me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

S&T : Science and technology

R&D : Research and development

PREST: Policy Research in Engineering, Science and Technology

SRI : Stanford Research Institute

IC : Information and Communication

ESD : Environment and Sustainable Development

ENR : Energy and Natural Resources

C&I : Construction and Infrastructure

M&M : Machine and Material

H&M : Health and Medicine

DASI : Defense, Aviation and Space Industry

A&F : Agriculture and Food

T&T : Transportation andTourism

Chem. : Chemistry

Tex. : Textile

CHAPTER 1

INTRODUCTION

At the end of the 20th century, we found ourselves in the midst of a technology driven rapid socio-economical change. Technology became part of our every day life. From our cell phones to laptops, from anti-allergic fabrics to transportation facilities, it became obvious that technological advances alter societies in every aspect. The importance of technological change as a source of economic growth and military power is accepted (especially in the post cold war period, technological innovation became an important determining factor for economic and military success of nations). Apart from military purposes, within a knowledge-based economy, industrial competitiveness became more and more dependent on new technologies and innovation capabilities of nations, regions, and firms.

Innovation occurs within an environment where social organizations, interaction and links among many actors are vital. Technological innovations introduce new products and processes or present improvements in the existing ones but the difficulty rests in the diffusion process of the newly achieved innovations, ideas. Innovation studies require a complex and multidisciplinary point of view in the sense that it not only requires the involvement of many diverse actors but also it underlines the importance of networks among these actors. Being a part of a network gains importance.

It is not possible to think about innovation and technological development separate from knowledge, knowledge production and learning processes. Knowledge is closely related to the innovation and therefore competitiveness of a production system. Competences in every level are mainly based on knowledge and established through learning processes. Campbell (1963) claims that knowledge-intensive societies and economies represent important clusters for the world system and their knowledge production trends define future evaluation of knowledge. Adding to this standpoint of placing knowledge base in the origin of innovation capability; this thesis will be based on the idea that common knowledge base identifies the connection between technologies and innovations.

There is usually a connection between technological changes in different areas; a change in a generic technology starts a chain reaction-like process by causing changes in other technology fields, merging some of them or by creating new technology areas. This kind of change in the generic technology usually paves the way for possible innovations and creates a line of change in the existing products or production capabilities in different sectors. New technologies, especially generic technologies, lead to new industry branches and create change in every field of economic activity (Göker, 1999).

Policy makers are paying special attention to what to support in this continuously changing, complex environment. In order to take accurate steps, they carry out programs that try to anticipate new technologies and their implications. Technological change imposes new opportunities and possibilities and as Perez (1986) states, it is important to be aware of the major elements of new techno-economic paradigm because only the predictions which could capture *the lines of force leading to structural change*, would be successful.

In broadest terms, reaching a list of generic technologies that can reinforce various areas of innovation takes place within the foresight process (Webster, 1999). Foresight offers a tool for getting the individual components of the national innovation system to communicate with each other, to discuss issues of longer-term common interest, to co-ordinate their respective strategies, in some cases to collaborate (Martin, 2002). In EU-supported FOREN program's report (FOREN, 2001) on regional foresight, it is stated that; "Foresight involves bringing together key agents of change and sources of knowledge, in order to develop *strategic visions* and *anticipatory intelligence*. Of equal importance, Foresight is often explicitly intended to establish *networks* of knowledgeable agents, who can respond better to policy and other challenges."

Delphi and scenario planning are very common foresight tools and sometimes used together to display a projection of future. Both of these techniques have emerged from military planning needs and they had been around for more than forty years now. Turkey implemented its first Delphi exercise in 2003 as a part of technology foresight study, with the aim of identifying strategic technologies to be supported for accomplishing the general aim of Vision 2023, welfare society.

The aim of the thesis is to investigate whether we can define technology clusters in terms of common expertise knowledge by applying methodology that Ronde (2001) suggested on Delphi results of Vision 2023, and to form future scenarios using these technology clusters.

In Chapter 2, a theoretical basis by examining innovation and technological systems mainly through the works of Schumpeter (1939), Perez (1983, 1986, 2004), and Carlsson and Eliasson (2001) will be established. In Chapter 3, a general overview of foresight and its difference from forecast will be presented together with explanation of Delphi and scenario building as future looking tools. The chapter will include Turkish Technology Foresight study of Vision 2023 and provide the results of Delphi survey conducted. Chapter 4 will investigate the techniques used in the thesis, look into cluster analysis, and explain hierarchical and non-hierarchical clustering techniques together with linear projection method. Chapter 5 will be based on the application of clustering methods on data set, provided by Delphi results of Vision 2023 and will try to achieve technology clusters and build future scenarios around them. Chapter 6 will conclude the thesis.

CHAPTER 2

THEORETICAL FRAMEWORK

In 1950s, dominating the economic thinking, traditional neo-classical economics influenced the general view of growth and technology-economy relation for a long time. The assumptions of the theory were modeling an unrealistic world developed on the ideal of a stationary state; a closed economy functioning under perfect competition with a production function of physical capital and labor. Technology was not accepted as an endogenous process determined within the system, rather it was determined outside the system; technology was given and it was a public good freely available to everyone. The theory failed to develop a satisfying explanation of economic growth; development was not a matter of discussion and the main aim was the analysis of a static equilibrium where the economy actually resided (Nelson & Winter 2002). The process of economic growth driven by technological change came into consideration by the works of Schumpeter.

2.1 Schumpeter and Notion of Innovation

Schumpeter set his agenda on how the economic change occurs as a result of the forces operating within the system. He was influenced from both theories; from the neo-classical equilibrium, the theory's micro-based approach on the one hand and Marxist dynamism, historical specificity with respect to technology, industry on the other hand (Fagerberg, 2003).

Distinct from the neo-classical theory, which relies on the ideal of static equilibrium, Schumpeter's theory was based on the idea that capitalist system was not stationary and that equilibrium of neo-classical theory was unattainable because of the constant disruption of qualitative change. Disequilibrium is inherent in the capitalist growth process and innovation, "setting up of a new production function", is the source of this dynamism (Schumpeter, 1939).

With his theory, Schumpeter included technology, which was accepted as given data until then, into the economics. Making a distinction between invention and innovation; he included invention in technical and scientific part of the event and innovation in the economical part.

> (As soon as it is divorced from invention) innovation is readily seen to be a distinct internal factor of change. It is an internal factor because the turning of existing factors of production to new uses is a purely economic process and, in capitalist society purely a matter of business behavior. (….) it is largely responsible for most of what we would at first sight attribute to other factors. [Schumpeter, 1939, p.86]

Not only in terms of mechanization, by developing new and more efficient machinery but also in terms of introducing new products, exploration of new markets and new organization types; innovation and technology are accepted as factors giving capitalism a dynamic and an evolutionary characteristic.

Although innovations designate social, technical and economical changes, not all the innovations bring out change with the same importance, domain and dimension. In that sense, it is useful to differentiate between *incremental* and *radical* innovations.

Perez (1986, 2004) defines *incremental innovations* as "successive improvements upon existing products and processes" and *radical innovations* as "the introduction of a truly new product or process". *Incremental innovations* occur in a routine manner, their outcomes may be important in the long-run in a collective manner but they aim and affect only the existing products and processes. They are the end-products of improvements. *Radical innovations* on the other hand, create major structural changes which require totally new skills or capital equipment. They do not emerge in the course of improvements of existing products or processes. According to Schumpeter (1939), these radical innovations are the products of entrepreneurs and are usually combinations of product, process and organizational innovation. It is entrepreneurs who convert the findings of scientists and inventors into new opportunities for growth in the economy.

Schumpeter (1939) offers innovations as the cause of cyclical performance of the capitalist economy. Schumpeter insisted that capitalist evolution represents a succession of "individual revolutions" and argued that important innovations do not occur randomly, but they tend to cluster in certain time-periods and sectors of the economy, and this likely gives rise to a discontinuous pattern of growth as *long waves*.

Perez (1983, 2004) takes this view and suggests that long waves are not only economic phenomenon but also they indicate the behavior of total socio-economic and institutional system, and they represent different *modes of development*, general patterns of growth based on a set of accepted social and institutional mechanisms. These patterns are determined by different and successive *technological styles* which can be described as the "best technological, productive common-sense".

Techno-economic paradigm changes or in other words *technological revolutions* are usually related with introduction of a *key factor* associated with a technology style. Key factor is the input which has low and descending relative cost, unlimited supply, massive and obvious all pervasiveness, and capacity to reduce the cost of capital, labor and products together with the capacity to change them qualitatively (Perez, 1983, 1986)

In Freeman, Clark and Soete (1982), the focus deliberately shifted from the dating of individual innovations to a system perspective in which the process of innovation-diffusion is studied as an interrelated whole. They suggested the term *technological system* for constellations of innovations which are technically and economically interrelated (Freeman, 1991). Such technological systems need not lead to long waves but may do so if a system is very large and of long duration or if the "bandwagons" of several different systems "roll" together.

2.1 Technological Systems

The idea that technologies' growing in systems rather than in isolation; necessitates the explanation of technological systems. Perez (1986) adopts Freeman's definition of *technological systems*, which may also be called *generic technologies* as "constellation of innovations, technically and economically interrelated and affecting several branches of production." These innovations lead to a move in general economic condition and some of which affect almost all other sectors, fuse with existing technologies, and provide new means of product and services.

According to Carlsson and Eliasson (2001), technological systems are not solely identified by technologies. There are three dimensions to technological systems namely; *cognitive, organizational and institutional, economic.*

Technology is the basis of technological systems and additionally, technology is determined by the knowledge base. In that sense, cognitive dimension – *design space* of a technological system is determined by the technological possibilities, clusters of technologies. Carlsson and Eliasson (2001) distinguish three modes of change in the design space: expansion of design space through the addition of new capabilities, integration of design spaces through co-evolution of its elements, accumulation of application-specific know-how.

Organizational and institutional dimension represents the interactions of the actors of technologies. It is the technological community that structures the design space and forms the social aspect of technological systems (Carlsson and Eliasson, 2001).

Related with organizational and institutional dimension, economic dimension of the technological systems can be defined as the actors converting technological possibilities into business opportunities (Carlsson and Eliasson, 2001).

Technological systems approach provides the suitable basis for studying the future of technologies and upcoming innovations in the sense that it combines technical, social and economical aspects of the technological change. As Perez (1986) states "the world of technically feasible is far greater than that of economically profitable and that of the socially acceptable".

CHAPTER 3

FORESIGHT AND TURKISH FORESIGHT CASE

The curiosity of people concerning what will happen in the next day, next year, next period of time in general, made them develop various ways to uncover the mysteries of future. People in China used some fortune sticks to find answers to their future related questions, Ancient Greeks turned to their oracle at Delphi to hear something about their future, some societies related the positioning of stars to the future and some societies tried to see it in the traces of coffee in a cup. By some means or other, activities concerning future, had always been in human being's today. Besides being in the center of strong interest, it had always been important to learn about the future in terms of making plans for surviving in the next days or next decades. Not only individuals but also states, organizations, industries and firms engaged in future related activities. They needed to learn about future in order to continue their existence. Mainly after Second World War, many forecasting methods were developed to discover, predict one possible future and what it would bring.

Together with the modern steps of scientific studies and technological advancements on the one hand, and insufficiency of future predicting tools on the other hand; not only the way of life, but also the understanding of future changed. Considering future as an inevitable outcome determined by destiny replaced with more sophisticated perception of it. Future is closer than ever, and we are aware of our capabilities in changing the course of time.

Especially with the arrival of 1980s, and failure of forecasting methods in foretelling the oil shock in 1973, activities on predicting that one possible future gave way to more complex future study activities; forecasting methods replaced with foresight, which accepts the fact that future holds many possible variations depending on the acts of people (Grupp and Linstone, 1999; McMaster, 1996). Human beings are active participants in the course of history and they shape it with their decisions and actions. Asking questions about future is important not only in terms of learning about the unknown anymore, but also in terms of achieving the power to manage it.

3.1 What is Foresight?

Foresight can be defined as 'the process involved in systematically attempting to look into the longer-term future of science, technology, the economy, the environment and society with the aim of identifying the emerging generic technologies and the underpinning areas of strategic research likely to yield the greatest economic and social benefits' (Cuhls, 2003). According to the Guide prepared by FOREN (2001), foresight is a 'systematic, participatory, future-intelligence-gathering and medium-long-term vision-building process aimed at present day decisions and mobilizing joint actions'. In a European Commission working paper (Von Schomberg, Pereira, and Funtowicz, 2005) foresight is roughly described as 'thinking, debating and shaping the future' and the document emphasizes its importance for science and technology development and improved quality of life in the knowledge society.

After Second World War, decision makers had to find a way to cope with uncertainty and applied forecasting methods. However, with the failure to expect the 1973 oil crisis, forecast activities lost popularity in the late 1970s and with the early 1990s, the interest shifted to foresight activities (Martin, 2002). Different from forecasting, foresight does not try to predict the future, rather it serves as a policy instrument as it tries to describe the range of possible futures for actors to prepare and shape the outcomes (Durand, 2003). Grupp and Linstone (1999) claim that traditional forecasting tools may be suitable for stable phases but they are insufficient in chaotic times and they see the emphasis of foresight on communication processes as a feature increasing the ability to respond. Table 3.1 summarizes the differences between forecasting and foresight.

Figure 3.1 illustrates the steps of foresight activities. Starting with the identification of different options for future, foresight selects the most desired one regarding certain criteria, like the meaning of that selection for today. The process contiues with the examination of the selection and making assumptions about that selection.

Then a decision is taken in accordance with the assumptions, which turns that option of future into a target (Cuhls, 2003).

As far as foresight, especially technology foresight, is regarded as a policy instrument, Cuhls (2003) sets some specific targets for foresight activities;

- To enlarge the choice of opportunities, to set priorities and to assess impacts
- To prospect for the impacts of current research and technology policy
- To ascertain new needs, new demands and new possibilities as well as new ideas
- To focus selectively on economic, technological, social and ecological areas as well as to start monitoring and detailed research in these fields
- To define desirable and undesirable futures
- To start and stimulate continuous discussion processes

Table 3.1 Foresight and Forecast Comparison (Source: Cuhls, 2003)

Foresight	Forecasting
• Basic points, needs, research questions are still open and looked for as part of the foresight process • More qualitative than quantitative • Looks for 'information' about the future for priority-setting • Brings people together for discussions about the future and networking, makes use of the distributed intelligence • Criteria for assessments and preparation for decisions • Communication about the future as an objective • Long-, medium- and short term orientation with implications for today • Finds out if there is consensus on themes • 'Experts' and other participants, very dependent on opinions.	• Basic points, topics and research questions have to be clarified in advance • More quantitative than qualitative • Questions what the future in the selected area might look like • More result oriented, can also be performed by individual people or single studies(depends on methodology) • Not necessarily assessments, different options and choices or the preparation for decisions • Describes future options, results more important than the communication aspects. • Long-, medium- and short term orientation as well as the path into the future are the major points • No information about consensus necessary • Mainly 'experts' and/or strict methodologies, less dependent on opinions

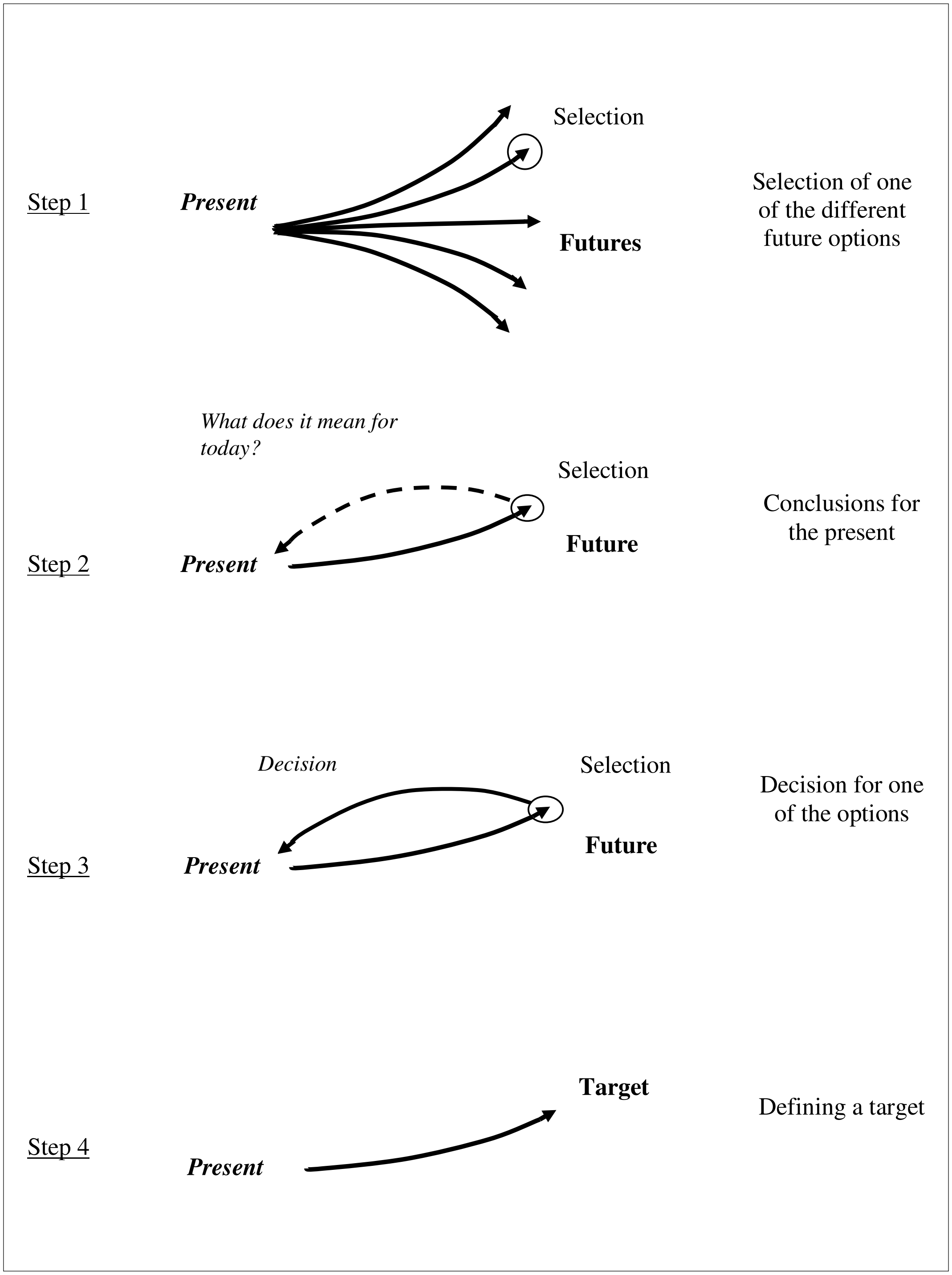

Figure 3.1 Steps in a foresight activity (Source: Cuhls, 2003)

The increasing popularity of foresight activities, especially technology foresight, can be based on the changing economical, social and political environment. Martin (2002) identified some driving factors of this change. One is the growing competition and change emphasizing the growing role of ‘new technology’ on the one hand and increasing importance of science and technology in terms of competence building in the other hand. Another factor is the limitations on public spending and governments’ necessity to identify the technologies to be funded. Increasing complexity resulting from closer interactions of research and technology, public and private sectors, different technologies and different producers of knowledge form another factor of change. Under these circumstances, foresight could carry out some required functions of providing an environment for making choices about science and technology and for identifying priorities; integration research opportunities with economic and social needs; stimulating communication and partnership among different actors of research, practitioners, users and funders (Martin and Johnston, 1999).

3.2 Foresight Techniques

The post-war nations of 1950s and 1960s, especially US, were highly concerned with military and defense related issues. It was a difficult problem to decide on which weapon system should be funded under uncertain conditions; uncertainty in terms of the success of chosen weapon systems, in terms of future political environment and in terms of other nations’ weapon systems (Bradfield et al., 2005). Under these circumstances, the US Department of Defense needed to satisfy two related needs; one is the need to obtain the consensus of expert opinion, which encouraged the development of Delphi technique and second is the need to develop models of future environments, which led to the use of scenario techniques (Bradfield et al., 2005).

3.2.1 Delphi Technique

The Delphi technique, taking its name from the ancient Greek oracle, is still the most used technique in foresight studies. The Delphi technique has mainly been used for evaluating technological development and its impacts but, its origins may be traced back to an experiment on the prediction of horse race results in 1948. In terms of its use in future studies, Delphi method, as many of future planning tools, emerged as a by-product of defense related research. It was first developed by the RAND Corporation in the 1950s as a method for gathering information about the future in an Air Force-sponsored study. The early studies at the RAND Corporation concerning the use of group information effectively showed that examining the character of the justifications given for short-term predictions could be used to evaluate the reliability of predictions and the background information had influence on the success of predictions (Dalkey, 1969). The set of procedures evolved from these findings and additional features received the name *Delphi*.

Pill (1971) defines the Delphi technique as “a method of combining the knowledge and abilities of a diverse group of experts to the task of quantifying variables which are either intangible or shrouded in uncertainty”. According to Linstone and Turoff (2002), Delphi is a “method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem”; Dalkey (1967, 1969) defines Delphi as “set of procedures for eliciting and refining the opinions of a group of people” and identifies three features;

- Anonymity; reducing the effect of individual dominancy
- Iteration and controlled feedback
- Statistical group response

Delphi is a structural survey, composed of two or more rounds. It mainly bases on interactive group communication and tries to achieve a common view of future, a consensus, generally in technology related issues among experts. Table 3.2 shows the comparison of Delphi and other group communication modes. The technique makes use of intuitive, tacit, explicit knowledge of the experts. De Loe (1995) assumes that a consensus of expert opinion has a better chance accuracy than the opinion of non-experts or individual experts. The reason is that they are aware of the limitations of current technology, have some idea of the direction of the change and feasibility of the possible alterations.

Table 3.2 Group Communication Techniques and Delphi (Source: Linstone and Turoff, 2002)

	Conference Telephone Call	Committee Meeting	Formal Conference or Seminar	Conventional Delphi	Real-Time Delphi
Effective Group Size	Small	Small to Medium	Small to Large	Small to Large	Small to Large
Occurrence of Interaction by Individual	Coincident with Group	Coincident with Group	Coincident with Group	Random	Random
Length of Interaction	Short	Medium to Long	Long	Short to Medium	Short
Number of Interactions	Multiple, as required by the group	Multiple, necessary time delays between	Single	Multiple, necessary time delays between	Multiple, as required by the individual
Normal Mode Range	Equality to chairman control (flexible)	Equality to chairman control (flexible)	Presentation (Directed)	Equality to monitor control (structured)	Equality to monitor control (structured)

Table 3.2 (continued) Group Communication Techniques and Delphi (Source: Linstone and Turoff, 2002)

	Conference Telephone Call	Committee Meeting	Formal Conference or Seminar	Conventional Delphi	Real-Time Delphi
Principal Costs	Communications	• Travel • Individual's Time	• Travel • Individuals Time • Fees	• Monitor Time • Clerical • Secretarial	• Communications • Computer Usage
	Time Urgent Considerations	Forced Delays		Forced Delays	Time-urgent Considerations
Other Characteristics	• Equal flow of information to and from all • Can maximize psychological effects		• Efficient flow of information from few to many	• Equal flow of information to and from all • Can maximize psychological effects • Can minimize time demanded of respondents and conferences	

There are at least two iterative rounds in a Delphi survey. The results of the previous round are given as feedback for reconsideration to the experts, who remain anonymous to each other through the survey. This feedback mechanism provides the experts with the opportunity to learn about the views of other experts and compare their answers with those of others. Anonymity makes sure that all participants get hearing and reduces the pressure to conform to the perceived views of senior colleagues (De Loe, 1995)

The survey composes of two sectors: topics and variables that resemble the things we want to learn about the topics like current situation, potential problems or feasibility of that topic. The experts are required to make estimations on variables of topics.

There exist two forms of Delphi; one is the most common paper-pencil version Delphi exercise, referred as *conventional Delphi*; second is the computer-based Delphi conference, referred as *real-tune Delphi*. Both versions undergo same four phases (Linstone and Turoff, 2002);

1) Exploration of the subject matter under discussion
2) Reaching an understanding of how the group views the issue
3) Evaluation of disagreements, finding out their reasons
4) Final evaluation

Delphi technique is often used for assessing technological development and its impacts but Linstone and Turoff (2002) suggest some conditions in which Delphi can be employed as an appropriate tool;

- If the problem is not suitable for precise analytical techniques but requires subjective judgments on a collective basis instead
- If the individuals whose contribution is required have diverse backgrounds and expertise but no communication history
- If the number of individuals whose contribution is required is not suitable for effective face-to-face interaction
- If frequent group meetings are infeasible in terms of time and cost
- If an additional group communication can increase the efficiency of meetings

- If strict disagreements among individuals necessitate a mediator or a mechanism for providing anonymity
- If conservation of the heterogeneity of the participants is important for validity of the results

In case one or more of the above-mentioned conditions occur, Delphi would be a desirable choice.

3.2.2 Scenario Techniques

Emergence of scenario techniques originates in the planning studies undertook by RAND Corporation for the US Airforce in 1950s. The US Military needed war game simulations and models of future environments in order to decide the weapon systems to be developed under the uncertain conditions of post-war period. Herman Kahn was then a military strategist at the RAND Corporation and he used scenarios to raise the awareness of the military by adopting 'thinking about the unthinkable' principle. After his departure from RAND Corporation, he founded the Hudson Institute where he started generating scenarios for public policy and social awareness.

Scenario is a tool for looking at the future; it provides possible or hypothetical pictures of future on narrative basis. Based upon certain information, it sets logical assumptions about what future might bring. Coates (2000), adopts the Webster's Ninth Collegiate Dictionary definition of scenario and takes scenarios to be "an imagined sequence of events, esp. any of several detailed plans or possibilities". Godet (1995) describes scenario as "a description of a future situation and the course of events which allows one to move forward from the original situation to the future situation". For Fahey and Randall (1998) the scenarios are "the projects of a potential future. They are a combination of estimations of what might happen and assumptions about what could happen, but they are not forecasts of what will happen".

Scenarios not only act as tools for future projection but also as testing ground for strategies. They are structured to explore and manage ambiguous situations and they provide a view of the future which conventional strategy analysis tools omitted, in other words they increase the relevance of a strategy where strategy aims creating competitive advantage (Perrottet, 1998).

Bradfield et al. (2005) agree that "there is a correlation between the adoption of scenario planning and environmental discontinuities". They are constructed after highly interactive, imaginative team processes (Schwartz and Ogilvy, 1998). Scenarios provide the managers and futurists with the environment for communicating on the sources and solutions of problems, challenge existing mental frontiers; assumptions, beliefs that affect perceptions.

Coates (2000) refers the value of scenario to its ability to convert complex elements into coherent, systematic, comprehensive, and plausible stories. In order to use scenario technique as an effective tool in decision-making processes or attracting public opinion, the scenarios have to be not illogical. This does not mean that scenarios must prove the future they projected will actually happen or offer proofs about their content; rather it means that this content should base on possible, credible and relevant hypothesis. According to Godet (1995), the hypotheses must be relevant, coherent and possible. Fahey and Randall (1998) claims that "plausible evidence should indicate that the projected narrative could take place (it is possible), demonstrate how it could take place (it is credible), and illustrates its implications for the organization (it is relevant)".

Fahey and Randall (1998), identifies the elements of a scenario. *Driving forces* which constitute the forces that shape the story; *logics* which explains the reasons of behavior of those forces; *plots* which describe how certain possible future might occur; and *end states* which define what possible futures look like (Figure 3.2).

3.2.2.1 Scenario Types

Scenarios can be categorized on different basis. The simplest categorization may be in terms of exploratory and normative scenarios. *Exploratory (future forward) scenarios* start from the past and present trends and project several scenarios, n*ormative (future backward) scenarios* on the other hand, start by selecting several futures and display the paths leading to them.

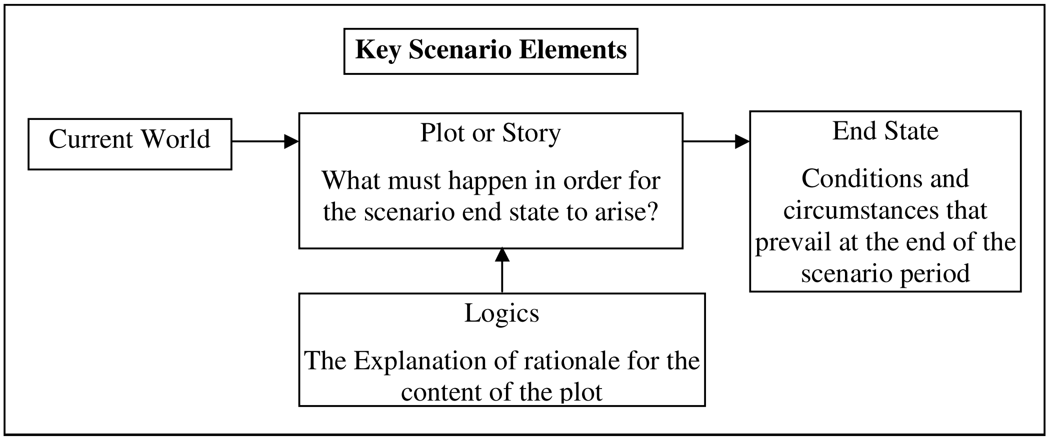

Figure 3.2 Key scenario elements (Source: Fahey and Randall, 1998)

Using more or less the same terminology, a scenario typology is presented in the work of Börjeson et al. (2006), according to which the main categories of scenarios are identified by their questions to the future. *Predictive scenarios* answer the question off '*what will happen?*' and they consist of two types; *forecasts*, responding the question '*what will happen on the condition that the likely development unfolds?*' and *what-if scenarios*, answering the question '*what will happen, on the condition of some specified events?*'. The second main category of scenarios is the *exploratory scenarios* those respond to the question '*what can happen?*' and sub-categorized into two types of *external* and *strategic scenarios*. External scenarios focus on the factors beyond the control of relevant actors and their question is '*what can happen to the development of external factors?*' while strategic scenarios respond to '*what can happen if we act in a certain way?*' to reveal the outcomes of strategic decisions. The last category comprises of normative scenarios with the question 'how can a specific target be reached?'. There are two types of normative scenarios; preserving scenarios of 'how can the target be reached by adjustments to current situation?' question and transforming scenarios answering 'how can the target be reached when the prevailing structure blocks change?'

Perrottet (1998) offers another categorization of corporate scenarios, providing possible corporate strategies against future threats and opportunities; scenarios for inspecting trends, identifying emerging opportunities; scenarios to prioritize large-scale research activities; scenarios for competitive gaming, aiming the information about the actions of the other side; and industry-focused scenarios.

Malaska (1995) mentions three different categories for scenarios adopted by companies; namely *issue scenarios*, *mission scenarios* and *action scenarios*. *Issue scenarios* provide general information and alternative futures for general issues, external to the company, without selecting a specific decision maker's point of view. Different from mission scenarios, they are publicly presented. *Mission scenarios* on the other hand, provide alternatives of where the company should or want to be in the future. *Action scenarios* combines the two and "provide answers to questions what the company can do taken the alternative futures of the issue scenarios given and the company's mission adopted".

A set of different scenario types were mentioned by Masini and Vasquez (2000): *utopian scenarios*, describing the best of the possible worlds, which is not achievable but useful for showing what is not possible; *trend scenarios*, showing the world in which everything stays the same as today, no change takes place and shows what will happen if things do not change; and *catastrophic scenarios* describing the worst of possible worlds.

3.2.2.2 Different Approaches

Scenarios were started to be used in different manners after Second World War in two different geographies. Future study tools in general were strategic tools in USA, in Europe on the other hand, they served as a tool to find the sources of and reconstruct strength. These geographies also formed two different centers for scenario techniques (Bradfield et al., 2005).

<u>Intuitive Logics (SRI-Shell Model)</u>

In the USA, when Khan founded the Hudson Institute after his departure from the RAND Corporation, he started using scenarios for public policy issues. Khan's methodology had already increased the interest in scenarios as a future study tool and the use of this methodology spread to private sector. Royal Dutch Shell and General Electric pioneered the use of scenarios in corporate planning but it was Royal Dutch Shell, which became famous.

Pierre Wack, a planner at Royal Dutch Shell, decided to use scenario technique as a tool for looking ahead to the year 1985 to investigate the future of the industry, as a part of 'Horizon Planning'. The resulting 'first generation scenarios' did not provide any hints about the future but proved to be useful in gaining a better understanding of the situation (Bradfield et al., 2005). They decided to keep on using scenarios for future planning which resulted in a success of identifying the coming oil scarcity followed by an increase in the oil prices.

In the meantime, Stanford University had established its own think-tank called Stanford Research Institute (SRI) and started to use scenarios as a long-range planning tool for political strategy, science, military and economics. They developed a methodology in parallel with the Royal Dutch Shell's scenario process; their method is known as the *intuitive logics* approach.

Intuitive logics, is the most common methodology for scenario planning. The methodology does not require any mathematical calculations and provides flexible and coherent scenarios. Masini & Vasquez (2000) and Wilson (1998) identify some characteristics of the approach, the most important ones:

- The method is decision-oriented
- Scenarios are frameworks for structuring management's perceptions about future environments
- In planning scenarios, causality rather than probability is taken into consideration
- Scenario planning is a process that draws on the knowledge and creativity of the participants

The Probabilistic Modified Trends

The probabilistic modification of extrapolated trends evolved out of the work of two former RAND alumnae; Olaf Helmer and Ted Gordon (Bradfield et al., 2005). They amalgamated two techniques, namely trend-impact analysis and cross-impact analysis.

Trend-impact analysis connects traditional forecasting techniques and quantitative factors. Statistical calculations and series of estimations are used for taking special factors, which might have an effect on the general trend, into consideration.

Cross-impact analysis on the other hand, includes the interdependency of the events, which were taken into account for scenarios. In the process, the probability of an event affecting the other is determined.

La Prospective

Although scenario techniques were started to be used at the same time with the US, French version, developed primarily by Gaston Berger, stemmed from different motives. Berger was more concerned with local issues of the long-term socio-political future of France. He adopted the methodology *La Prospective* with the aim of creating "an acceptable scenario-based methodology for developing positive images or normative scenarios of the future and to lead these images into the political arena where they could serve as a guiding vision to policy maker" (Bradfield et al., 2005).

In 1970s, Michel Godet began to develop scenarios basing on the *perspective* of Berger; he used largely mathematical, computer-based probabilistic method to compute cross-impacts (Bradfield et al., 2005; Ringland, 1998).

Bradfield et al.(2005) introduce a comparison of Probabilistic Modified Trends, La Prospective and intuitive logics approaches in terms of objectives, methodologies, orientations and importance attributed to the experts, which is summarized in Table 3.3. It can be seen that La Prospective and Probabilistic Modified Trends share many things in common, mostly because they both use mathematical models in visioning the future.

Table 3.3 Comparison of Scenario Techniques (Source: Bradfield et al., 2005)

	Innovative-Logics Models	*La Prospective* Models	Probabilistic Modified Trend Models
Purpose of the scenario work:	Multiple; from a once-off activity making sense of situations and developing strategy, to an ongoing activity associated with anticipation and adaptive organizational learning.	Usually a once-off activity associated with developing more effective policy and strategic decision and tactical plans of action.	A once-off activity to enhance extrapolative prediction and policy evaluation.
Scenario perspective:	Descriptive or normative	Usually descriptive, can be normative	Descriptive
Scope of scenario exercise	Can be either broad or narrow scope ranging.	Generally a narrow scope.	Narrow scope focused on the probability and impact of specific events on historic trends.
Scenario horizon year: Methodological orientation:	Varies: 3-20 years, process orientation.	Varies: 3-20 years, outcome orientation.	Varies: 3-20 years, outcome orientation.
Nature of scenario team participants:	Internal – scenarios developed by a facilitated from within the organization.	Combination of some key individuals from within the organization led by an expert external consultant.	External – scenario exercise undertaken by expert external consultant.
Tools commonly used:	Generic – brainstorming, STEEP analysis, clustering, matrices, system dynamics and stake-holder analysis	Proprietary – structural (Micmac) and actor (Mactor) analysis, morphological analysis, Delphi, SMIC Prob-Expert Multipol and Multicriteria evaluation	Proprietary Trends Impact and Cross Impact Analysis, Monte Carlo simulations
Scenario starting point:	A particular management decision, issue or area of general concern.	A specific phenomenon of concern.	Decisions/issues for which detailed and reliable time series data exists.

3.3 Turkish Technology Foresight and Vision 2023

'Turkish Science Policy: 1983-2003' was the first detailed science and technology policy document for Turkey, emphasizing the role of technology for development and suggesting priority areas of technology, first attempt to define critical technologies in Turkey. Although this document could not be implemented, it resulted in the establishment of a new institution, Supreme Council for Science and Technology, as the highest science and technology policy making body of Turkey in 1983.

In 1993, in its second meeting, Supreme Council for Science and Technology approved the document entitled 'Turkish Science and Technology Policy: 1993-2003'. The document was the recognition of the turning point from 'science policies' to 'science and technology policies' approach and its reflection in science and technology policy of Turkey from 'building a modern research and development infrastructure' to 'innovation oriented' national policies. It accepted a national science and technology policy for the next ten years, stressing the role of science and technology in national economy, sustaining economic growth, improving the living standards, and in international competition. In 1995, the document has been rebuilt as 'The Project for Impetus in Science and Technology', and formed the science and technology chapter of the Seventh Five Year Development Plan.

Yet, the lack of shared science and technology vision and commitment to the policies among governments, universities, and public and private sectors led to the failure in target achievement. Under these circumstances, in its sixth meeting, Supreme Council for Science and Technology decided in the preparation of Turkish science and technology strategies document which would aim at formulating new national science and technology policies and priority areas for achieving an innovative economy and society in 2023. The project entitled "Vision 2023: Science and Technology Strategies" has been approved by the Council in its seventh meeting in December 2001.

3.3.1 Vision 2023

The raison d'être of Vision 2023 project is the "creation of welfare society which masters science and technology, able to produce newer technologies and capable of converting technological developments into social and economic benefits".

The project was planned to include the studies of;

- Determination of current situation of Turkey in terms of science and technology;
- Determination of long term science and technology trends in the world,
- Determination of science and technology demand of Turkey, regarding 2023 targets,
- Determination of vital strategic technologies which would provide the achievement of these targets,
- Suggestion of policies which would help to develop and/or obtain these technologies.

Vision 2023 project involved four sub-projects, which were; 'National Technological Capacity, Technology Inventory', 'Research and Development Manpower', and 'Research and Development Infrastructure' collecting and evaluating data on the current science, technology, and innovation capacity of the country; together with the first national foresight exercise of Turkey 'Technology Foresight' sub-project.

Technology Foresight Project constituted the core of Vision 2023. The project adopted socioeconomic and thematic panels and a Delphi survey as tools of foresight. Figure 3.3 illustrates the project process.

The results of this project together with the results of other three sub-projects determining the existing .science-technology situation and of the future potential of Turkey provided the basis for, '2003-2023 Strategy Document' (Figure 3.4). This document included; the science and technology vision of Turkey, R&D priorities and strategic technologies of Turkey, policy recommendations necessary to create capability in these priority areas.

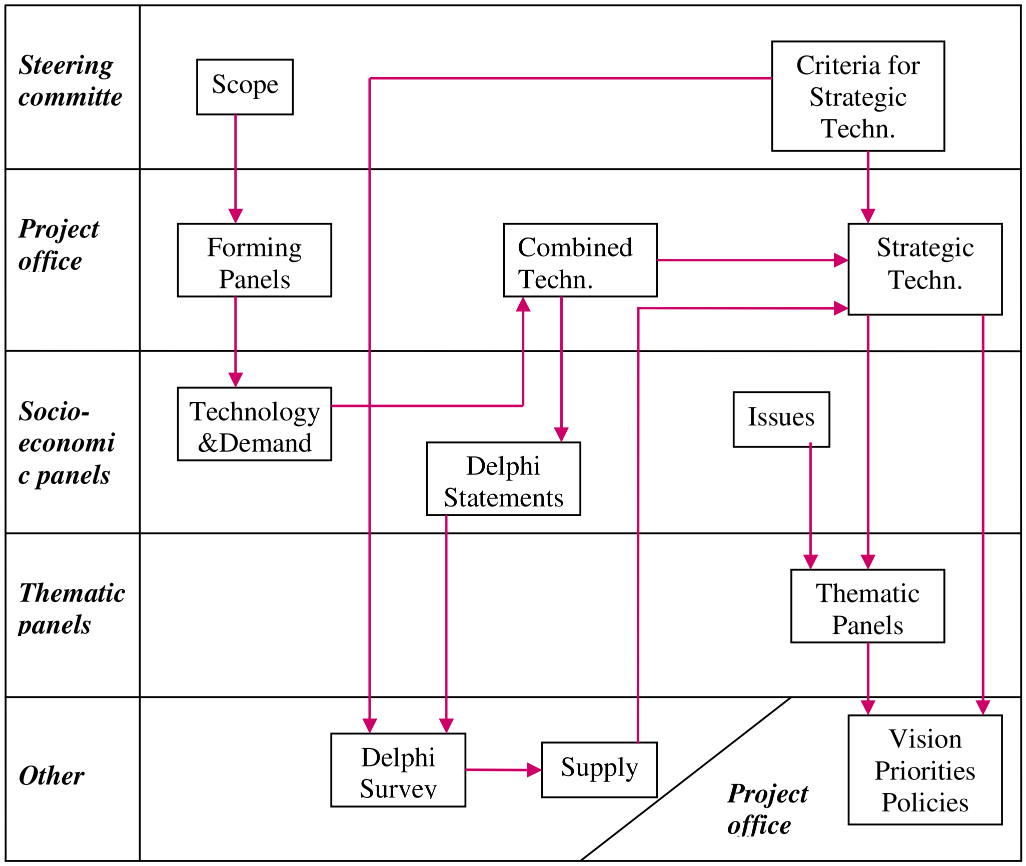

Figure 3.3 Process of Technology Foresight Project (Source: Tümer, 2002)

The main idea of technology foresight study in Vision 2023, was formulating the science and technology policy for 2003-2023 and besides, utilization of process benefits. Focusing on science and technology, the study was expected to draw a picture of strategic technologies and research and development priorities for Turkey, raise the awareness on science and technology, and satisfy the effective and wide-range participation.

Ten panels were formed on certain socioeconomic fields and two others on crosscutting issues of education and human resources, and environment and sustainable development, the subjects were underlined by socioeconomic panels.

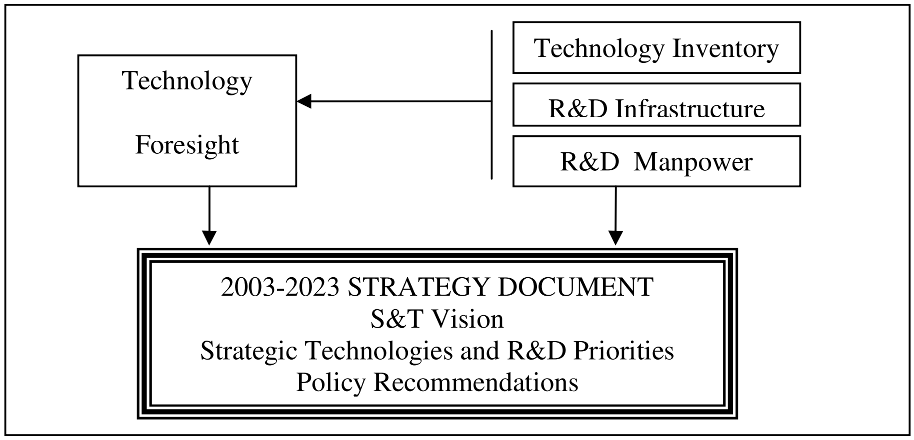

Figure 3.4 Road to Strategy Document (Source: Tümer, 2002)

These panels were;

1. Education and Human Resources (Thematic panel)
2. Environment and Sustainable Development (Thematic panel)
3. Information and Communication
4. Energy and Natural Resources
5. Health and Pharmaceuticals
6. Defense, Aeronautics and Space Industries
7. Agriculture and Food
8. Machinery and Materials
9. Transportation and Tourism
10. Textiles
11. Chemicals
12. Construction and Infrastructure

The socioeconomic panels studied on;

- Turkey's vision for 2023,
- Socio-economic objectives to be achieved in order to realize this vision,
- Technological activity areas necessary to reach these objectives,
- Technology fields and capabilities needed for these topics,

- Turkey's current standing (SWOT analysis)
- Reports

Panels were carried out in a four period process;

1- Development of vision (July 2002- January 2003; preliminary report: 24.01.03)
2- Wider consultation (February 2003- March 2003; revised report: 28.03.03)
3- Delphi surveys (March 2003- June 2003)
4- Policy recommendations (May 2003- July 2003; final report: 24.07.03)

Around 200 panel meetings and workshops took place between July 3, 2002 and January 24, 2003. With the completion of preliminary panel reports on January 28, 2003, the concerned groups were invited for their criticisms and contributions and each panel publicized their initial works among the related actors.

All panels, except the Education and Human Resources Panel, prepared more than 1200 statements that are important in realizing their 2023 visions. The final list included 413 statements forming around 20 questions.

The two-rounded Delphi surveys were expected to question;

- The feasibility of technological capabilities foreseen by the panels both for Turkey and world,
- The effect of these technological changes on the determined criteria for strategic technologies.

Each of the ten socioeconomic panels reviewed the results of the two round Delphi survey and completed their final reports, which were submitted on July 24, 2003. Each panel also submitted technological roadmaps for achieving the foreseen technological developments in each 'Technological Activity Area' in their final reports.

Among the technology fields foreseen by panels and evaluated by experts in Delphi surveys, technologies with high ranks in feasibility index and importance index were determined as strategic technologies. The feasibility index was determined by using panel and Delphi results. On the other hand, importance index was determined by using Strategic Technology Criteria.

Strategic technology criteria were determined by the prioritization of steering committee baring the results of the panels in mind. They were simply the answers given to the questions ‘what are Turkey’s priority expectations from science and technology in terms of 2023 targets?’ and ‘what features of a specific technology makes it preferable to others?’ The sample answers were competitiveness, productivity, or energy saving, etc.

Strategic technologies, in which Turkey has to develop or gain capability, were grouped in 8 main headings as

1. Information Technologies
2. Biotechnology and Gene Technologies
3. Energy and Environmental Technologies
4. Materials
5. Mechatronics
6. Nanotechnology
7. Design Technologies
8. Production Methods and Machinery

Strategy groups were established under these eight technology fields. The task of strategy groups were to prepare “technology roadmaps” for those specific technology fields they assume as priority fields, and of critical importance for the realization of the technology activity areas proposed by the panels. For each of the groups, experts of the relevant technologies have identified the strategically important technology fields and policies and strategies necessary to gain capability on these fields. Finally, the 20-year Science and Technology Strategy Document based on the findings and recommendations of sub-projects and highlighting the technology roadmaps on priority technology fields, was prepared and submitted to the Supreme Council, and subsequently to the Government for consideration and adoption, in September 2004.

3.3.2 Delphi Results

The variables, according to which the topics (technologies) were examined, were concluded after a joint workshop of technology panels, project office and a specialist from PREST (Policy Research in Engineering, Science and Technology). Project office determined six main variables, in line with the specific needs of the country. Priority and feasibility variables have also been derived using these variables. This way, from Delphi application, results regarding current situation, initial capability, required initial science and technology capability, time of realization, science and technology policy tools, their contribution to benefit of Turkey, and priority and feasibility variables have been obtained.

1.CURRENT SITUATION

It is important to have information on the current situations of sectors / processes which might develop Delphi statements in the future. That is why respondents have been asked questions relating to the current situation in Turkey for each Delphi statement.

Considering all main and sub-variables in the survey, expert views asses the situation more positively, compared to all other respondents.

Current situation information has been examined in terms of researcher potential, R&D infrastructure, masters in relevant fundamental sciences, innovative levels of firms, and existence of competitive firms. For each of these sub variables, *None*, *Poor*, *Sufficient* and *Strong* choices have been offered; and the respondents were asked to choose one of these.

Figure 3.5 illustrates the axis that was created to simplify the evaluation. Considering the situation in terms of responds to all Delphi statements, we see that only the sub variable “masters in relevant fundamental sciences” have received responses which fall within the “sufficient” interval of the axis, from both respondents (1,65) and experts (1,90). For “researcher potential” sub-variable, all respondents are within the “poor” interval of the axis (1,46) while experts are positioned in “sufficient” interval

(1,71). All respondents (1,22) and experts (1,41) agree on the weakness of our R&D infrastructure. Responds to "existence of competitive firms" is 0,98 for all respondents, and 1,11 for experts.

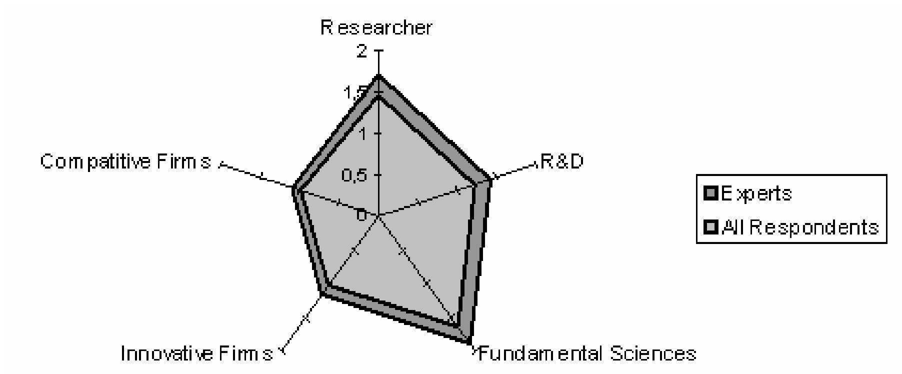

Figure 3.5 Delphi results for "Current Situation" (Source: TUBITAK, http://vizyon2023.tubitak.gov.tr/yayinlar/delfi/)

<u>Researcher Potential</u>

Experts have assessed the existing researcher potential in health and drugs, sustainable development; and agriculture and food areas as *sufficient*. While existing researcher potential in machinery and material, and information and communication is considered sufficient, the assessments of these areas take place in the negative end of the axis.

<u>R&D Infrastructure</u>

Experts' assessment of R&D infrastructures of agriculture and food, and transportation and tourism is "sufficient" but with levels much above the average. However, even these fields are close to "poor" interval. The fields, which were considered as "poor" with assessment levels below the average, are machinery and material, environment and sustainable development, information and communication, textile and defense, aeronautics and space.

Masters in relevant fundamental sciences

Experts have assessed the current situation of mastery in fundamental sciences, in chemistry, health and drugs, energy and natural sources; and environment and sustainable development as positive with levels much above the average. Though positive, the assessment that relatively has the lowest levels is in machinery and material; and information and communication areas.

Existence of Competitive Firms

Experts consider the existence of competitive firms in the current situation as "poor". Nonetheless, experts in textile, and transportation and tourism assess the current situation as "poor" but close to "sufficient". Moreover, experts in chemistry, environment and sustainable development, energy and natural sources, machinery and material, defense, aeronautics and space, health and drugs consider the existence of competitive firms under current situation to be below the average and "poor".

The Innovative Capability of Firms

Experts in textile area find the innovative capability of firms "sufficient"; however, the tendency in other areas is towards negative. Experts in transportation and tourism, chemistry, energy and natural sources, construction and infrastructure, health and drugs, information and communication, defense, aeronautics and space, machinery and materials considered this capability to be "poor".

2.INITIAL CAPABILITY

In order for realizing each Delfi statement Turkey may need to obtain one or more initial capability target. It is aimed to determine to which area these capabilities fall; basic research, applied research, industrial development or competitive industrial development. Considering the general means, both respondents and experts are included in industrial and applied research slice. In industrial and applied researches, the highest rate of expert group is for transportation and tourism. Experts in information and communication; construction and infrastructure; defense, space and aeronautics underline the high levels of applied and industrial development. According to the experts, the areas that need industrial development capability least are agriculture and food, chemistry, and environment and sustainable development.

3. TIME OF REALIZATION

In case consistent and stable strategies are applied in our country, the time of realization for each Delfi statement is required. The periods presented to the respondents are: 2003-2007, 2008-2012, 2013-2022, 2023 and after, never.

The choices of experts represent a more rapid realization compared to respondents. In case of implementation of consistent strategies, it is seen that a central tendency that all Delfi statements will have been realized in the year 2013 emerges. 38,9% of experts, and 38,6% of respondents point that period. The second period that has been chosen is the first period (2003-2007) with rates of 33% for experts, and 27,5% for respondents.

Assessing the periods in terms of activity areas, experts chose the first period only for environment and sustainable development, construction and infrastructure, and textile while the agreement on the other areas fall into the second period (Table 3.4).

4.POLICY TOOLS

In this major variable of the survey, the effectiveness of policy tools in terms of aimed Delfi statements is questioned:

- R&D infrastructure support
- R&D project support
- Initialization support
- Guided projects
- Human resources support
- Public procurement support

Table 3.4 Delphi results for the"Time of Realization" (Source: TUBITAK, http://vizyon2023.tubitak.gov.tr/yayinlar/delfi/)

Fields (Expert opinions unless otherwise is written)	The time gap when the statement occurs	Main tendency (%)	The ratio of the other four periods (%)
Information and Communication	2008 – 2012	42	58
Environment and Sustainable Development	2003 – 2007	40,1	59,9
Energy and Natural Resources	2008 – 2012	39,1	60,9
Construction and Infrastructure	2003 – 2007	39,8	60,2
Construction and Infrastructure (All)	2008 – 2012	36,2	63,8
Chemistry	2008 – 2012	41,1	58,9
Machine and Material	2008 – 2012	37,7	62,3
Health and Medicine	2008 – 2012	38,8	61,2
Defense, Aviation and Space Industry	2008 – 2012	39,9	60,1
Agriculture and Food	2008 – 2012	39,2	60,8
Textiles	2003 – 2007	41,4	58,6
Textiles (All)	2008 – 2012	38,1	61,9
Transportation and Tourism	2008 – 2012	46,3	53,7

R&D infrastructure support represent the supports regarding the infrastructure investments such as purchase of equipment required in the R&D activities of the concerned institutes, which are not directly included in the project. R&D project support is the financial aids, which will be provided within an R&D project. Initiation support includes supports such as seed money given for prompting the establishment and improvement of firms open to technologic improvement. Guided projects are R&D projects, which are managed by research supporting institutions through other institutions or consortiums, in a specific field with determined topic and limiting conditions. Human resources are policy tools for training researchers who are qualified for R&D activities, at home and abroad; and bringing qualified local/foreign researchers in Turkey. Public supply indicates the policies aiming for the scientific and technological improvement of a country through determining the R&D and local contribution levels of firms in purchase of specific products/services required by the public.

Examining the required policy tools shows that both respondents and experts highlight R&D project support (31.5%). This is followed by guided projects (20.2%). 15.8% of experts demand R&D infrastructure support; 14.6% demand initiation support; 9.0% human resources support; and 6% demand public supply support.

5.CONTRIBUTION TO TURKEY

It is important to define the contribution of each Delphi statement to Turkey, in terms of some particular criteria. The following criteria have been questioned for each Delphi statement:

- *Science, Technology, and Innovation Capability:* Turkey's increasing its capabilities in science, technology, and innovation processes.
- *National Value Added:* Increasing domestic production and assessing domestic resources in this direction
- *Life Standard*: Increasing lifetime, quality and comfort throughout the country
- *Competition Power:* Turkey's increasing its share in national and international markets, especially in goods and services with high added value, through a raise in productivity
- *Environmental Sensitivity and Energy Efficiency:* Evaluating environmental problems, protecting the environment, and increasing energy savings and efficiency

In order to understand general tendencies regarding the main variables, an axis has been created. On this axis, average points between 0-0.49 indicate "no contribution", 0.5-1.49 interval indicates "little contribution", 1.5-2.49 interval indicates "contributes" and 2.5-3.0 interval indicates "major contribution" preferences.

Examining the responds in this regard, we see that all questioned statements fall within "contributes" (1.5-2.49) interval. The tendency of experts towards this reply is heavier than those of respondents.

Science, technology, and innovation capability: 2.19

In terms of contribution to science, technology, and innovation; the highest rank is occupied by defense, aeronautics, and space; and the lowest rank is occupied by information and communication; environment and sustainable development; and transportation and tourism.

National Added Value: 2.17

Agriculture and food; energy and natural sources; defense, aeronautics and space; and health and drugs take place in the foremost levels of the axis in terms of *national added value*. On experts' account, textile; information and communication; and chemistry occupy the last ranks of the axis in terms of this criterion.

Quality of Life: 2.13

Considering contribution to *life standard* criterion, it appears that environment and sustainable development has the highest levels, and that the position of this field on the axis is the one, which approximates the "major contribution" interval most. According to experts, in terms of this criterion, several fields are above the average. These fields are transportation and tourism; health and drugs; construction and infrastructure; information and communication; energy and natural sources; and agriculture and food.

Competitiveness: 2.0

For experts, textile, agriculture and food, and machinery and material fields occupy the first three ranks in terms of contribution to competitive power. The lowest assessment on this criterion belongs to construction and infrastructure, energy and natural sources, an environment and sustainable development.

Environmental sensitivity and energy efficiency: 1.92

Fields which are assessed by the experts to be above the average are environment and sustainable energy, energy and natural sources, agriculture and food, transportation and tourism, construction and infrastructure, chemistry, and machinery and material, and the experts placed health and drugs; defense, aeronautics and space; and information and communication into the last three ranks.

a) Importance

In order to determine the important technologies for Turkey, Steering Committee has created a composite index, and determined the criteria as well as their weights. Considering the experts, the importance index ranges between 84.92 and 41.08; and this indicates a relative “optimistic” assessment interval even though there are no big differences between the intervals. According to the experts, the fact that seven of statements regarding different fields have fallen within the first twenty is worth noting. In statements related to one same field, textile is in the first 20 ranks with four. From health and drugs; and transportation and tourism, two for each statement have entered the list. It is realized that transportation and tourism are effective fields in mixed statements.

b) Feasibility

In determining the feasibility index, the main variables of *current situation* from the previous sections of the report, and required *initiation capability* have been used. Indices corresponding to the experts’ feasibility assessments range between 68.60 and 21.79. Experts have the most optimistic view compared to all respondents. Of the statements, which were considered the most feasible ones, six are from textile, three from transportation and tourism fields. Information and communication, and health and drugs take place in the first 20 with two statements for each. There are five statements, which take part in both construction and infrastructure, and transportation and tourism (i.e. mixed statements). In addition, three statements from transportation and tourism directly take place within the first 20.

3.3.3 Science and Technology Strategy

Recognizing the importance of science and technology for the social and economical development of societies in the long term; Supreme Council for Science and

Technology decided to prepare National Science and Technology Policies: 2003-2023 Strategy Document in its sixth meeting. The document (TUBITAK, 2004) would guide the development and emphasize a common shared vision for Turkey in 2023;

- Being motivated to establish a fair and long lasting peaceful environment and a peaceful world,
- Having a democratic and righteous legal system,
- Having citizens who decide for the future of their country,
- Ensuring that health, educational and cultural needs of its citizens will be met by the State,
- Seeking for sustainable development; have an equal income distribution,
- Reaching a competence level in science, technology and innovation; be a productive country which can increase its net value added with its own brain power.

After determining where you want to be in the future, it is important to resolve the instruments, which will take you there. With the knowledge of science and technology's power and importance in shaping societies, Technology Foresight Panels concluded four socio-economic targets and necessary competency areas in accordance for the realization of this 2023 Vision of Turkey (TUBITAK, 2004). The socio-economic targets and needed competencies can be stated as:

1. Gaining *competitive advantage* and raising the share in international markets. In order to achieve this goal, competency should be built on;
   - Flexible production/flexible automation process and technologies
   - Development of information-intensive products with high value added; and becoming a global design and production center for consumption goods
   - Hygienic production skills
   - Agricultural production
   - Development of space and defense technologies and material technologies

2. Increasing *life quality* of our people; and in order to achieve this goal, essential competencies are;

- Food safety and reliability
- Health and life sciences
- Healthy and modern urbanization and the required infrastructure
- Modern and safe transportation systems

3. Achieving *sustainable development*; and in order to achieve this goal, competency should be built on energy technologies, environmental technologies, and evaluation of natural resources.
4. Strengthening *information and communication infrastructure*; in order to keep up with the transformation occurring in a world in which societies' capabilities *knowledge production*, and convert that knowledge into economic and social utility shapes national economies and social life.

With the purpose of achieving these socio-economic targets and required competencies, priority technological activity areas and eight underlying strategic technologies were determined with the study of panels. The criteria in defining these strategic technologies are:

- Having high demand range and depth in other sectors as well as its own sector,
- Having the potential of human capital and knowledge accumulation,
- Having a potential to meet strategic requirements and a high possibility of contributing in the technological and economic development of the country in the present and the foreseen future,
- Would grant our country's place in global markets or will facilitate Turkey's holding and strengthening its current place.

The strategic technologies and relevant technological activity areas with priority for the next twenty years can be listed as follows:

➢ Information and Communication Technologies

Information and communication technologies have played an enormous role in shaping the social, economical and technological structures of societies in the last quarter of 20th century. The technology foresight study showed that this role of information and communication technologies would continue to exist for the following years and not only information and communication bases of societies but also other technology fields would be nourished from it. The most strategic of these information and communication technologies for Turkey's 2023 vision are;

- Integrated Circuit Technologies Design And Production
- Display Unit Production Technologies
- Wide Band Technologies
- Production Technologies for Sensors for Vision Perception

➢ Biotechnology and Gene Technologies

Developed countries have succeeded in converting the scientific developments in the field of molecular biology, cell biology and genome studies into economic value. These developments not only stayed as life quality-increasing achievements but also became driving forces of economies. With the guidance of strategic road maps, the document (TUBITAK, 2004) states that Turkey will have the chance to become a global force in these technologies within a short period of twenty years, which would create an economic and technologic advantage. The most strategic of these technologies for Turkey's 2023 vision are;

- High Scale Platform Technologies: Structural Genoscience, Functional Genoscience, Transcripteomics, Proteomics, Metabolomics
- Recombinant DNA Technologies
- Molecular Diagnosis And Treatment: Cells And Stem Cells
- Therapeutic Protein Drug Production And Controlled Release Systems
- Bio-Informatics

➢ Mechatronics

Mechatronics and bio-mechatronic system applications are very important for civil and military service sectors in some of which Turkey has international competitive advantage and aims to increase that power. Both the target and necessity to achieve competitive advantage in international markets, and to increase life standard levels make mechatronics a strategic technology for Turkey. In accordance, Turkey should focus on the following technology areas in order to be the producer and exporter of mechatronics and human integrated mechatronic systems;

- Micro/Nano Electro-Mechanical Systems (MEMS/NEMS)
- Robotics And Automation Technologies

➢ Nanotechnology

Nanotechnology would generate an improvement in welfare level, in economy and national security of the countries within a decade, so it is important to play an effective role in developing nanotechnology. The following areas seem to have strategic importance for Turkey to produce and store the scientific, technologic and industrial knowledge;

- Nanophotonics Nonoelectronics Nanomagnetism
- Nanomaterial
- Nanocharacterization
- Nanofabrication
- Nano-Scale Quantum Information Processing
- Nano-Biotechnology

➢ Production Process and Technologies

Turkey has developed a certain power in the automotive, home comfort and home electronics, textile production activities in the last century and had the benefits of employment and production utilities. However, a strategy which would increase and sustain the current advantage on the one hand and which would include investments in the emerging technologies on the other hand should be followed. The relevant strategic technology fields are;

- Flexible And Extreme Production Techniques
- Rapid Prototyping Technologies
- Face/Interface, Thin Film And Vacuum Technologies

- Metal Shaping Technologies
- Plastic Piece Production Technologies
- Wielding Technologies
- Machining Technologies

➢ Material Technologies

With the function of supplying inputs to all kinds of economic activities, material sector forms one of the main technologic areas, which will change the characteristics of industrial production. The emergence of the advanced materials, which will be used in defense, aeronautics, microelectronics, communication and automotive sectors, improves in accordance with the transformation of material science to a multidisciplinary, process-weighted area; advanced material field will create important attraction centers. In the extent of materials, Turkey should focus on;

- Boron Technologies
- Composite Material Technologies
- Polymer Technologies
- Smart Material Technologies
- Magnetic Electronic And Opto-Electronic Material Technologies
- Low And High Strength Material Technologies

➢ Energy and Environmental Technologies

It is inevitable for Turkey to build competency in the technologies, which would maximize the benefits from our energy resources and provide competitive advantage. To the document (TUBITAK, 2004) technologies that will allow managing our water resources in a sustainable way, and preventing pollution and recovery of waste should be focused on.

- Hydrogen Technologies And Fuel Cells
- Renewable Energy Technologies
- Energy Storage Technologies And Power Electronics
- Nuclear Energy Technologies

- Environmental Friendly And High Efficiency Fuel And Burning Technologies
- Water Refinement Technologies and Waste Treatment Technologies

➢ Design Technologies

Among production activities, design has a fundamental and major role in developing new products and technologies, and innovations to increase the benefit of current products and technologies. The following technology areas should be focused on;

- Virtual Reality SW And Virtual Prototyping
- Simulation And Modeling Software
- Grid Technologies And Parallel And Distributed Computation Software

In order to gain capability in science technology and innovation the strategy document proposes;

- To focus on the strategic technology fields and supporting scientific research fields.
- To allocate resource for research and development
- To train necessary man power and allocate necessary resource
- Political
- To create awareness within social strata
- To observe the realization process of Vision 2023 and calculate its outcomes and to establish a sustainable system for evaluation
- To establish a system for periodical reevaluation of Vision 2023 under changing socio-economical conditions

CHAPTER 4

CLUSTER ANALYSIS

4.1 What is Cluster Analysis?

Cluster analysis covers a range of techniques used for finding homogenous groups of objects in a data set. Objects of the same group are similar to each other, while the objects of different groups are as different as possible from each other in terms of the variables considered. Everitt (1974) made a set of definitions of a cluster:

- "A cluster is a set of entities which are alike, while entities from different clusters are not alike."
- "A cluster is an aggregation of points in the test space such that the distance between any two points in the cluster is less than the distance between any point in the cluster and any point not in it."
- "Clusters may be described as connected regions of a multi-dimensional space containing a relatively high density of points, separated from other such regions by a region containing a relatively low density of points."

In cluster analysis, there are no predetermined classes or groups. Jain and Dubes (1988) state that distinct from discriminant analysis and pattern recognition, there are not category labels in cluster analysis as it tries to achieve right organization of data not to attain rules of future data classification.

Unlike model-driven multivariate data analysis, reasoning is data-driven in exploratory data analysis (Backer, 1995). Statistical tests might be used for verification of a model or hypothesis but, as an exploratory tool, cluster analysis does not seek to confirm a predetermined hypothesis in the first place. Janowitz (2002) identifies the goal of cluster analysis as "to find some evidence of some inner structure for the data", in order to discover what data tell us, it is possible to try several models on the same data. On the other hand, deciding which clustering method to use depends on the type of data too. In that sense it is important to mention data types.

4.1.1 Type and Scale of Data

Data can be *binary*, *discrete* or *continuous*. *Binary* data have two possible values. Kaufman and Rousseeuw (2005) talk about two kinds of binary data; *symmetric binary data* where values have equally same weight, and *asymmetric binary data* where values are not equally important. Binary data is often coded as zero and one, where zero usually means the absence of a certain feature and one indicates its presence. *Discrete* data have finite and small number of values but no intermediate values where *continuous* data may have any real value in a fixed range of values. Number of visitors a patient had is discrete data as there cannot be 5.2 visitors, whereas room temperature in different times forms continuous data.

Data can be measured on different scales, on *ratio-scale*, *interval-scale*, *ordinal-scale* or *nominal-scale*. Data scale shows the meaning and importance of numbers. On *nominal* scale, numbers are quantitatively meaningless as they are only used in stead of names. On *ordinal* scale, numbers have meaning in terms of ranking. On *interval* scale, interpretation of the numbers depends on a measurement unit. Numbers have absolute meaning on *ratio* scale as an absolute zero exists together with a unit of measurement.

4.1.2 Pattern Matrix & Proximity Matrix

For clustering techniques to be employed, data have to be represented in the form of *pattern matrix* or *proximity matrix*.

Pattern matrix: $n \; X \; d$ is the pattern matrix which comprises the set of representation of n objects by d measurements. Each row of this matrix identifies a pattern while each column indicates a *feature*, or measurement.

Proximity matrix: Before applying any method for clustering, a measure of distance between objects, index of proximity between them has to be defined. This proximity index can be derived either from the pattern matrix or from the raw data and a proximity matrix gathers the indices of proximity between pairs of objects in a matrix which each row and column represent a pattern (Jain and Dubes, 1988; Backer, 1995).

A proximity matrix can either be in terms of *similarity* or in terms of *dissimilarity* among objects. Janowitz (2002) provides a commonly used setting for defining similarity and dissimilarity measures;

Objects, $S = \{s_1, s_2, \dots, s_n\}$.

Mapping $p: S \, X \, S \rightarrow R$ such that

$$p\ (s_i, s_j) \geq 0$$

$$p\ (s_i, s_j) = p\ (s_j, s_i)$$

$$p\ (s_i, s_i) = p\ (s_h, s_h)$$

$$p\ (s_i, s_j) \leq p\ (s_h, s_h) \text{ for all } i, j, h$$

Higher values of p represent more similar pair of objects where $p\ (s_i, s_i)$ is the highest value.

Mapping $d: S \, X \, S \rightarrow R$ such that

$$d\ (s_i, s_j) \geq 0$$

$$d\ (s_i, s_j) = 0$$

$$d\ (s_i, s_j) = d\ (s_j, s_i)$$

Higher values of d represent more dissimilar, less similar pair of objects.

When p is a similarity measure, a dissimilarity coefficient d is defined by

$$d\ (s_i, s_j) = p\ (s_k, s_k) - p\ (s_i, s_j)$$

4.1.3 Proximity Indices

Depending on the type and scale of data, different proximity indices are used. For binary and nominal-scaled data, *matching coefficients* are used and for continuous and ratio-scaled data, *Minkowski distances* are suitable.

Matching coefficients measure similarity, while Minkowski distances measure dissimilarity. In Figure 4.1, an example of binary pattern matrix is given.

Pattern Matrix

0	1	0	0	1	1
0	1	1	1	1	0
0	0	1	0	1	0
0	0	1	1	1	1
0	0	0	0	1	1
0	0	1	1	0	1
0	0	0	1	0	1
0	0	0	1	1	1
0	0	1	1	0	0
0	1	1	1	0	1

Corresponding dissimilarity matrix

Simple Matching Coefficient

0	0.50	0.50	0.50	0.17	0.67	0.50	0.33	0.83	0.50
	0	0.33	0.33	0.67	0.50	0.67	0.50	0.33	0.33
		0	0.33	0.33	0.50	0.67	0.50	0.33	0.67
			0	0.33	0.17	0.33	0.17	0.33	0.33
				0	0.50	0.33	0.17	0.67	0.67
					0	0.17	0.33	0.17	0.17
						0	0.17	0.33	0.33
							0	0.50	0.50
								0	0.33
									0

Jaccard Coefficient

0	0.60	0.75	0.60	0.33	0.80	0.75	0.50	1.00	0.60
	0	0.50	0.40	0.80	0.60	0.80	0.60	0.50	0.40
		0	0.50	0.67	0.75	1.00	0.75	0.67	0.80
			0	0.50	0.25	0.50	0.25	0.50	0.40
				0	0.75	0.67	0.33	1.00	0.80
					0	0.33	0.50	0.33	0.25
						0	0.33	0.67	0.50
							0	0.75	0.60
								0	0.50
									0

Russel-Rao Coefficient

0	0.67	0.83	0.67	0.67	0.83	0.83	0.67	1.00	0.67
	0	0.67	0.50	0.83	0.67	0.83	0.67	0.67	0.50
		0	0.67	0.83	0.83	1.00	0.83	0.83	0.83
			0	0.67	0.50	0.67	0.50	0.67	0.50
				0	0.83	0.83	0.67	1.00	0.83
					0	0.67	0.67	0.67	0.50
						0	0.67	0.83	0.67
							0	0.83	0.67
								0	0.67
									0

Figure 4.1 An example of the use of Matching Coefficients in the case of binary data (Source: Backer, 1995)

Minkowski distance is the generalized distance function (Jain and Dubes, 1988). For two points *i* and *j*, whose distance is measured on *k* dimensions (variables), the Minkowski distance would be;

$$d\,(i,j) = (\,|x_{i1} - x_{j1}|^p + |\,x_{i2} - x_{j2}|^p + \ldots + |\,x_{ik} - x_{jk}|^p\,)^{1/p}$$

When p=1, Minkowski distance is *Manhattan distance*; when p=2, Minkowski distance is *Euclidean distance*. In general, varying *p,* changes the weight given to the larger and smaller distances (Johnson and Wichern, 1998). These most common Minkowski distances are;

Euclidean distance, defined by

$$d\,(i,j) = (\,|x_{i1} - x_{j1}|^2 + |\,x_{i2} - x_{j2}|^2 + \ldots + |\,x_{ik} - x_{jk}|^2\,)^{1/2}$$

Manhattan distance, defined by

$$d\,(i,j) = (\,|x_{i1} - x_{j1}| + |\,x_{i2} - x_{j2}| + \ldots + |\,x_{ik} - x_{jk}|\,)$$

For computing similarity between two objects i and j, we define *a*, as the number of variables that equal 1 for both objects; *b* as the number of variables *h* for which x_{ih} = 1 and x_{jh} = 0; c as the number of variables *h* for which x_{ih} = 0 and x_{jh} = 1; and finally d, as the number of variables that equal 0 for both objects. A relevant *association table* would be as follows;

		Object *j*		
		1	0	
Object *i*	1	*a*	*b*	*a* + *b*
	0	*c*	*d*	*c* + *d*
		a + *c*	*b* + *d*	

Figure 4.2 Association table (Source: Kaufman and Rousseeuw, 2005)

In accordance with these associations, the most common matching coefficients can be defined;

Simple matching coefficient,

$$s(i,j) = (a+d)/(a+b+c+d)$$

Jaccard's coefficient,

$$s(i,j) = a/(a+b+c)$$

Russel-Rao coefficient,

$$s(i,j) = a/(a+b+c+d)$$

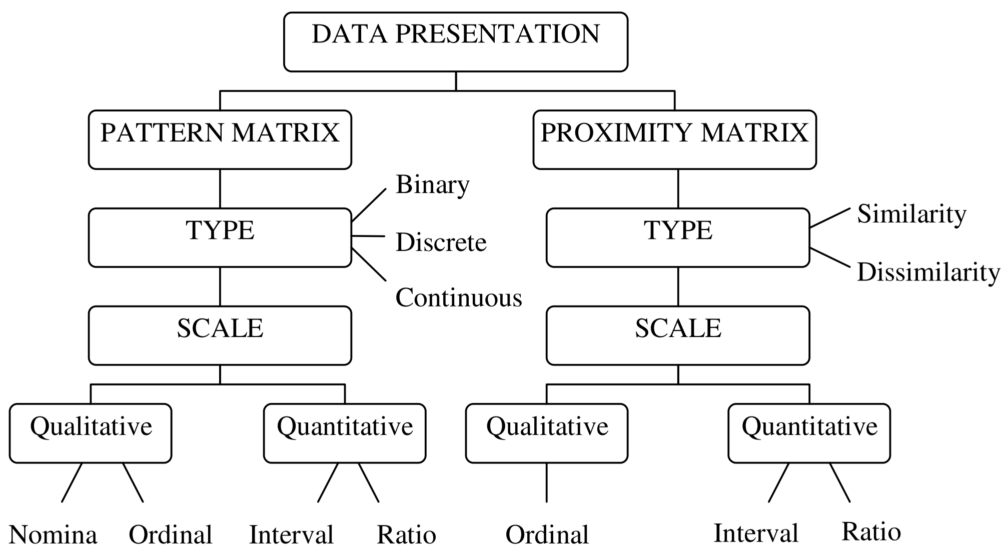

Figure 4.3 Formats, types and scales for data (Source: Jain and Dubes, 1988)

It is important for our study to note that proximity indices can also be binary, discrete or continuous. Figure 4.3 illustrates various types and scales of data. Coefficients differ in terms of dealing with different types of data, especially binary data for this study. As Kaufman and Rousseeuw states (2005), for symmetric data, it is natural to work with *invariant* similarities, the most known of which is simple matching coefficient; and for asymmetric binary data, with *non-invariant* coefficients like Jaccard's coefficient. The simple matching coefficient weights the matches of 0's the same as the matches of 1's, but Jaccard's coefficient ignores the matches of 0's.

4.2 Clustering Methods

4.2.1 Non-hierarchical (Partitional) Clustering

Partitional clustering tries to achieve predetermined number of clusters out of data and as Kaufman and Rousseeuw (2005) state, each group must contain at least one object while each object must belong to exactly one group. Classification of objects into k groups, where k is fixed and given by the user, is an attempt to uncover 'natural' groups present in the data.

Different from hierarchical clustering methods, non-hierarchical clustering methods generate a single partition of the data. Hierarchical methods generally use the proximity matrix whereas partitional methods make use of the data in the form of a pattern matrix (Jain and Dubes, 1988).

A non-hierarchical clustering method should be able to make sure that a pattern in a cluster is more similar to other patterns in the same cluster than to the patterns in another cluster. A clustering criterion must be adopted. Most partitional clustering methods base on *square error criterion*. That means, the method would try to achieve the partition which would minimize the square error, also called the *within-cluster variation*, for a fixed number of clusters. This variation tends to decrease as the number of clusters increases. There are different criteria for obtaining a partition and each criterion imposes a certain structure on data.

K-means clustering is the most popular non-hierarchical clustering method. The method works by assigning each item (point) to the cluster whose centroid is the nearest. Johnson and Wichern (1998) mention two different starts for non-hierarchical methods and refer to MacQueen for the process of K-means clustering. Non-hierarchical methods in general and K-means clustering method in particular, can start from either an initial partition of items into k clusters or arrangement of k initial centroids (seed points) for the clusters. Then the clustering proceeds by assigning items to the clusters with the nearest centroid. Centroids would be recalculated each time a cluster receives an item or loses one. The process ends when there are no more reassignments.

Partitional methods' speed and capacity to compress large datasets make them useful especially for engineering applications. On the other hand, while trying to adjust an outlier data to *k* elliptical clusters, the method distorts natural formation. Especially forcing the binary data into *k* groups would lead to nonsensical clusters (Johnson, and Wichern 1998). Therefore, K-means clustering methodology is not used in this study.

4.2.3 Hierarchical Clustering

Backer (1995) defines hierarchical clustering as a sequence of partitions in which each partition is nested into the next partition in the sequence. There are two major types of hierarchical clustering; *agglomerative* or *divisive*. *Agglomerative methods* start with each object being a cluster itself and continue with progressively combining the two nearest clusters until all objects form a single cluster; it is the series of successive fusions of *n* objects into groups.

Divisive methods work in reverse, starting with *n* objects being in a single cluster, they proceed by separating these clusters into successively smaller clusters until each object becomes a cluster itself. Both divisive and agglomerative methods can be displayed in the form of a *dendrogram*, a two-dimensional tree diagram which shows the merges or divisions made at each successive level (Johnson and Wichern, 1998). Cutting a dendrogram at any level, defines a clustering and identifies clusters.

There are different agglomerative hierarchical clustering techniques among which the linkage techniques and Ward’s method are the most common ones. Differences between techniques grow out of the way they define distance between clusters. Figure 4.4 illustrates the difference between different techniques.

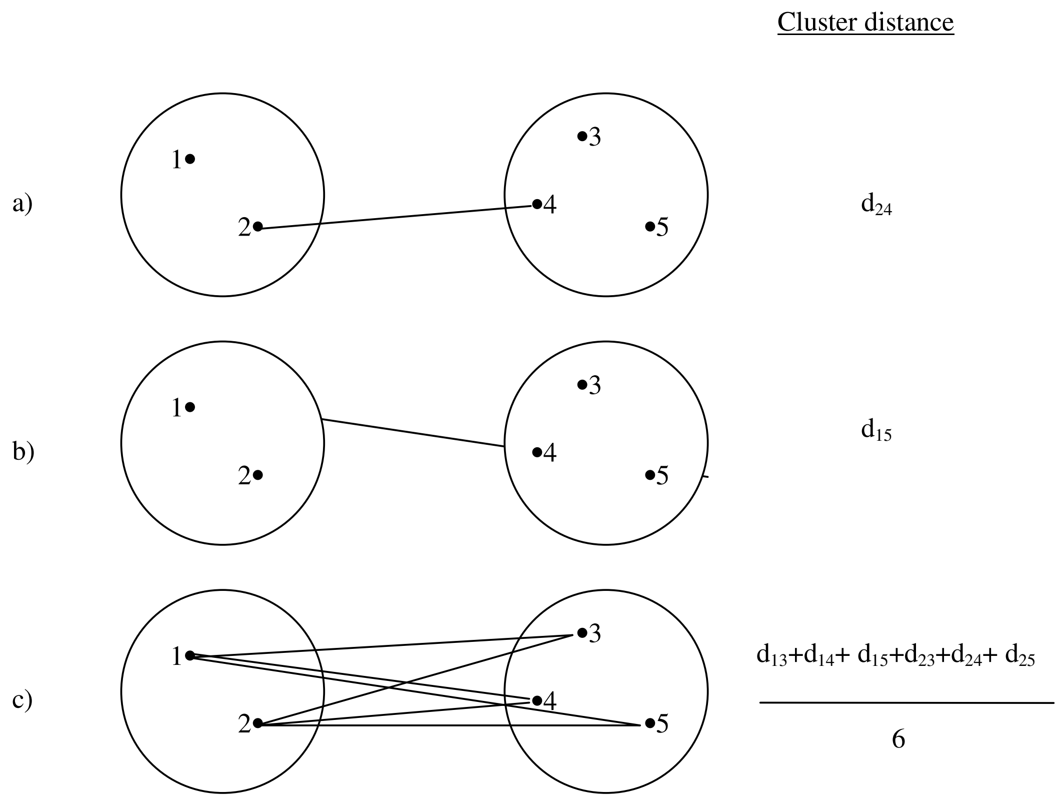

Figure 4.4 Inter-cluster distance for a) single linkage, b) complete linkage, c) average linkage (Source: Johnson & Wichern, 1998)

4.2.3.1 Single-Linkage Clustering (Nearest Neighbor)

In *single-linkage clustering*, distance between clusters is defined as the distance between the closest pair of objects, each from distinct clusters. At each stage of clustering, the clusters with minimum distance are merged.

Let Cluster A = $\{a_1, a_2, \dots, a_m\}$,

Cluster B = $\{b_1, b_2, \dots, b_n\}$

Then the *single linkage distance* between A and B is;

$$min\ (d(a_i, b_j)|\ i = 1,2, \dots, m; j = 1,2, \dots, n)$$

4.2.3.2 Complete-Linkage Clustering (Farthest Neighbor)

In *complete-linkage clustering*, distance between clusters is defined as the distance between the farthest pair of objects, each from distinct clusters. The distance between two clusters is given by the value of longest link between the clusters.

Let Cluster A = $\{a_1, a_2, \dots, a_m\}$,

Cluster B = $\{b_1, b_2, \dots, b_n\}$

Then the *complete linkage distance* between A and B is;

$$max\ (d(a_i, b_j)|\ i = 1,2, \dots, m; j = 1,2, \dots, n)$$

4.2.3.3 Average Clustering

In *average clustering*, distance between clusters is defined as the average distance between all pairs of individuals within each other.

Let Cluster A = $\{a_1, a_2, \dots, a_m\}$,

Cluster B = $\{b_1, b_2, \dots, b_n\}$

Then the *average distance* between A and B is;

$$sum(d(ai,bj)|\ i = 1,2, \dots, m; j = 1,2, \dots, n)/(mn)$$

4.2.3.4 Ward's Method

This method uses variance analysis to evaluate the distances between clusters. The clusters are generated by minimizing the *within-cluster variance*. According to Johnson and Wichern (1998) Ward's method bases hierarchical clustering on minimizing the loss of information growing out of joining two clusters. The loss of information is taken to be the total sum of squared deviations of every point from the mean of the cluster to which it belongs; in other words, it is an increase in the error sum of squares criterion.

The methods are useful for finding different types of clusters. Nearest neighbor method works well when clusters come too close to each other and elongated or chain-like clusters are to be found; furthest neighbor method works well when the plotted clusters form distinct bunches, it tends to create compact clusters; group average method finds ball-shaped clusters; and Ward's method tends to create clusters of small size. Figure 4.5 illustrates different types of clusters and appropriate methods in accordance.

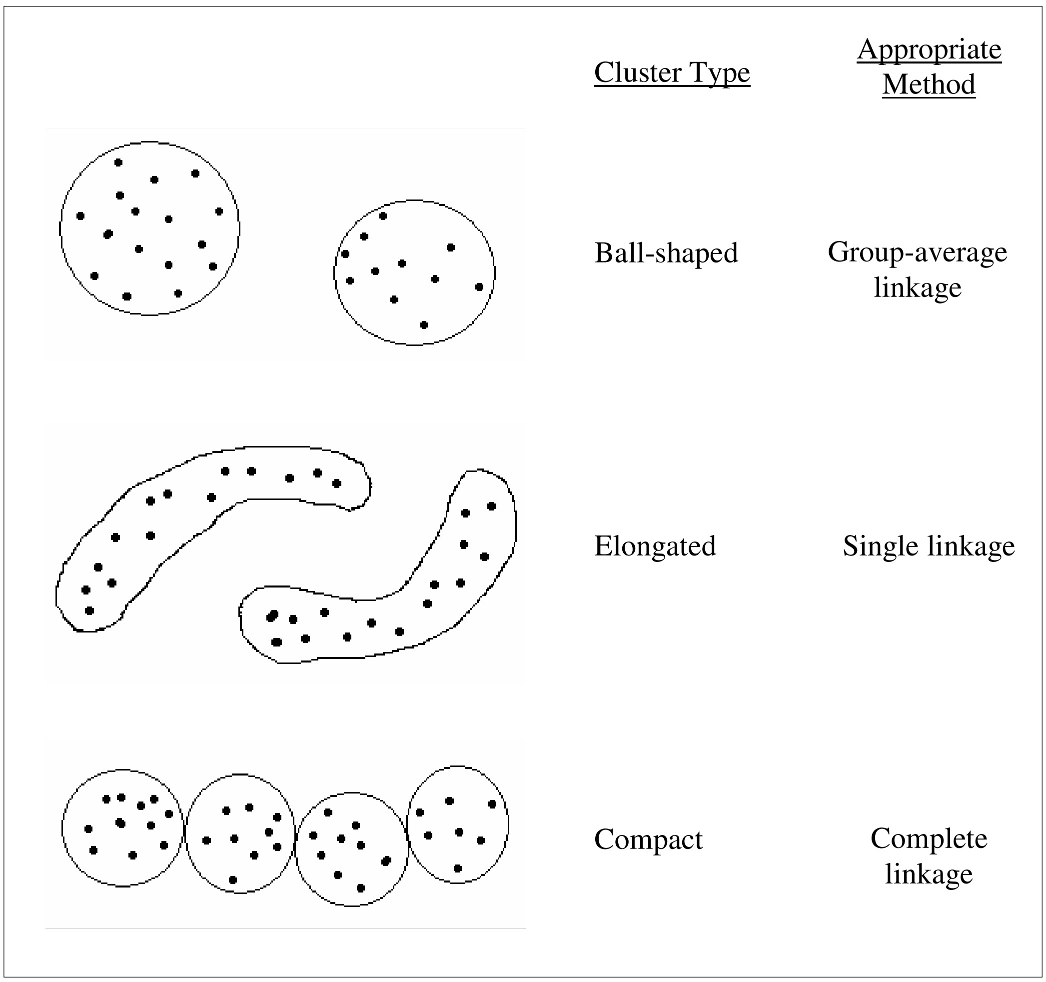

Figure 4.5 Some types of clusters and clustering methods for finding them (Source: Kaufman and Rousseeuw, 2005 with slight adjustments).

4.3 Projection Methods

Projection methods are comprised of a set of techniques that map a set of *n* *d*-dimensional patterns onto an *m*-dimensional space (Backer, 1995). Since the main objective for studying projection methods in cluster analysis is to permit the visual examination of multivariate data, *m* is generally selected as 2 or 3.

If a projection method employs linear combination of the original *d* feature vectors to express the *m* new features, it is called *linear projection* (Jain, Dubes, 1988).

Linear projection can be defined as;

$$y_i = \boldsymbol{H} x_i, \qquad for\ i=1,2,\ldots,n$$

Here, y_i is an *m*-place column vector, x_i is a *d*-place column vector and H is an m x d matrix.

Eigenvector projection is one of the well-understood and commonly used linear projection methods, which is also known as *Karhunen-Loeve method* or *principal component analysis*. The eigenvectors of the proximity matrix *R* of normalized pattern matrix define a linear projection that replaces features in the raw data with uncorrelated features. Since *R* is a *d x d* positive definite matrix, all of its eigenvalues are real and can be labeled such that;

$$\lambda_1 \geq \lambda 2 \geq \lambda_3 \geq \ldots . \geq \lambda_d$$

The set of corresponding eigenvectors (*principal components*) defines the projection matrix into the new coordinate system where each axis (or feature) is uncorrelated. Therefore, **H** matrix can be defined as

$$\mathbf{H} = [c_1\ c_2\ c_3 \ldots . c_m]$$

where c_i is the eigenvector corresponding to the eigenvalue λ_i

An appropriate selection of *m* should be done such that maximum variance is retained in the new *m* dimensional space. Since the eigenvalues of the proximity matrix R are the sample variances of the new space (Jain and Dubes, 1988), percentage of the variance retained can be calculated as,

$$r_m = 100 \text{ x } (\lambda_{1+} \lambda_{2+} \lambda_{3+\ldots} + \lambda_m)/(\lambda_1 + \lambda_{2+} \lambda_{3+\ldots+} \lambda_d)$$

As an example, Figure 4.6 shows a raw data set with two dimensions.

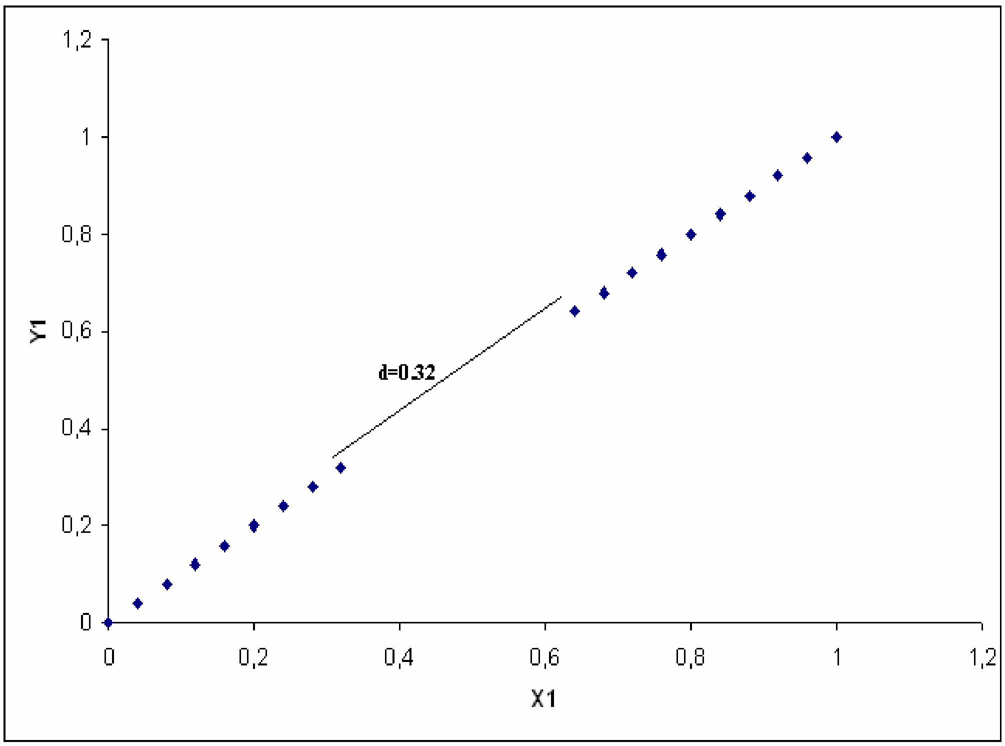

Figure 4.6 An exemplary raw data set

If the dimensionality of the data reduced to one without changing the coordinate axes, the representation is as shown in Figure 4.7. Note that the distance between the points changes through the projection. The eigenvalue projection method rotates the coordinate axes such that new axes represent the data better in reduced dimensionality.

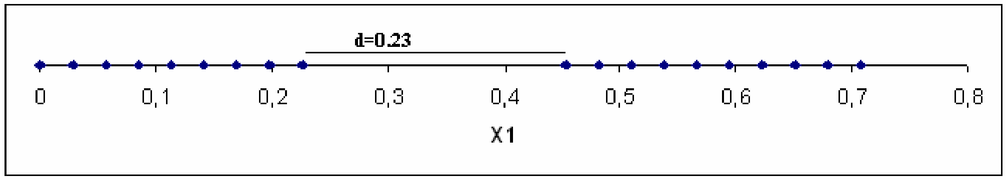

Figure 4.7 Projection of raw data set of Figure 4.6 onto X1 axis

In this case, as shown in Figure 4.8, reselection of coordinate axes as *X1'* and *X2'* does not change the nature of the data. However, representation of the raw data using only *X1'* as principal component contains much more information of the distribution of the points relative to each other, which is illustrated at Figure 4.9.

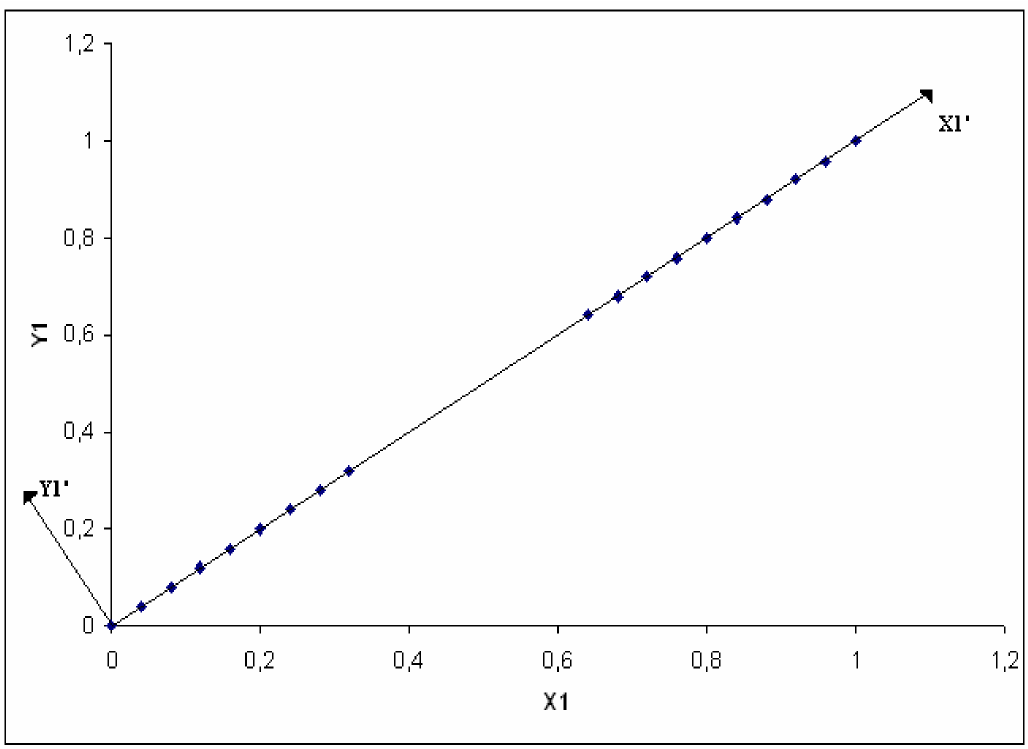

Figure 4.8 New coordinate system, X1' and Y1'

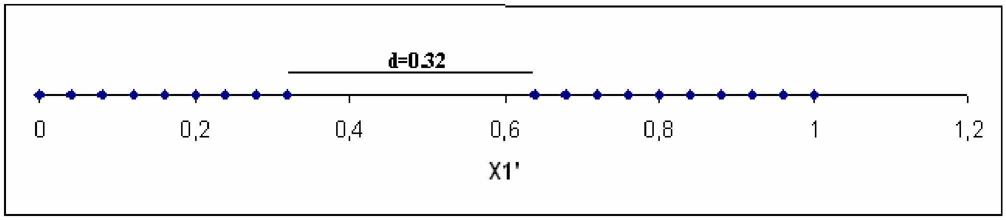

Figure 4.9 Projection of raw data onto X1' (data represented in one dimension without any information loss)

For this specific case, by selecting the new coordinate system, we can represent the data in one dimension without loosing anything

CHAPTER 5

FINDING TECHNOLOGY CLUSTERS

5.1 Data

Delphi investigation in Turkish technology foresight project, Vision 2023, obtained from TUBTAK, constitutes the raw data for this study. In technology foresight project, two different methods were used together to identify technology demands and supply. The methodology involved mainly socioeconomic and thematic panels on the one hand and a Delphi survey on the other hand. After around 200 panel meetings and workshops, 1200 statements that are important in realizing 2023 visions were prepared, with the exception of Education and Human Resources Panel. The final list included 413 distinct statements. Some Delphi statements were in relation to more than one technology area so they were used more than once, the outcome is a total number of 517 Delphi statements (Table 5.1).

Delphi method, as it was applied in Vision 2023, had some differences from other Delphi applications around the world. Different from general practice, instead of a separate commission, Delphi applications were carried out by technology panels; both paper-based survey sheets and virtual survey sheets were used and this use of virtual survey made it possible to reach out more experts, not only the ones whom the paper-based questionnaire forms were mailed. Experts not only answered the statements of their expertise but also they had the chance to answer the statements they see relative to their subject matter (TUBİTAK, 2004).

Survey sheets were mailed to a list of experts who were chosen from technology foresight panels and TUBİTAK database and 1636 of them answered the statements. Other than the ones from the list, 687 more experts responded the survey through internet, the use of virtual survey sheets. A total number of 2323 experts joined Vision 2023 Delphi survey (Table 5.2).

Table 5.1 Distribution of Statements to Panels and Answering Ratios (Source: TUBITAK, http://vizyon2023.tubitak.gov.tr/yayinlar/delfi/)

Name of Panel	Number of Statements	Answer per Statement	
		Partially Completed	Fully Completed
Information and Communication	58	146	134
Environment and Sustainable Development	27	135	125
Energy and Natural Resources	47	113	105
Construction and Infrastructure	29	88	82
Chemistry	52	113	106
Machine and Material	67	101	94
Health and Medicine	46	127	118
Defense, Aviation and Space Industry	45	88	82
Agriculture and Food	73	188	178
Textiles	40	47	44
Transportation and Tourism	33	98	92
Total	413	188	111

Table 5.2 Participation to Delphi Survey (Source: TUBITAK, http://vizyon2023.tubitak.gov.tr/yayinlar/delfi/)

Respondents	Number	Percentage
Number of experts on the list	7016	100
Respondents from the list	1636	23,3
Respondents that are not on the list	687	9,8
Total respondents	2323	33,1

The respondents were required to determine their own expertise level by choosing one of the levels; (1) No expertise, (2) A little expertise, (3) Enough expertise, (4) High expertise. Table 5.3 illustrates the distribution of expertise among the respondents, and Table 5.4 illustrates the distribution of experts among panels. According to TUBİTAK (2004) data, 34.4% of the participants identified themselves as “experts”, those who chose levels of “enough expertise” (3) or “high expertise” (4).

Table 5.3 Distribution of Respondents’ Expertise Level (Source: TUBITAK, http://vizyon2023.tubitak.gov.tr/yayinlar/delfi/)

	Expertise Level			
	None	Poor	Decent	Fair
All Fields	28,1	37,4	25,5	8,9
Information and Communication	18,2	39,8	31,8	10,2
Environment and Sustainable Development	26,9	37,4	25,8	9,9
Energy and Natural Resources	28,5	37,1	26,6	7,8
Construction and Infrastructure	35,5	36,4	21,5	6,6
Chemistry	33,5	36,1	23,9	6,5
Machine and Material	28,3	41,0	24,4	6,2
Health and Medicine	30,6	36,3	23,4	9,7
Defense, Aviation and Space Industry	21,1	41,5	28,4	8,9
Agriculture and Food	30,7	34,8	23,4	11,0
Textiles	21,7	36,9	29,8	11,7
Transportation and Tourism	35,3	36,2	22,2	6,3

5.2 Hypothesis

The study intends to reach meaningful groupings of Delphi statements in the first place and generate technology clusters from them by supposing that the expertise levels of respondents represent the knowledge-base of the relevant statements, technologies. Following Ronde (2001), it is assumed that if an expert evaluates himself as having ‘high’ or ‘enough’ expertise level for two different topics, it means that these two topics are “close in terms of science and/or technological base”.

Table 5.4 Distribution of Respondents to Panels (Source: TUBITAK, http://vizyon2023.tubitak.gov.tr/yayinlar/delfi/)

Name of Panel	Number of Statements	All Respondents	Experts
Information and Communication	58	143,2	60,2
Environment and Sustainable Development	27	128,7	48,1
Energy and Natural Resources	47	109,8	38,1
Construction and Infrastructure	29	84,6	24,2
Chemistry	52	111,7	34,1
Machine and Material	67	97,0	29,9
Health and Medicine	46	124,8	41,4
Defense, Aviation and Space Industry	45	85,7	32,3
Agriculture and Food	73	186,8	64,6
Textiles	40	46,1	19,1
Transportation and Tourism	33	94,2	27,1
Total	413	107,4	38,3

5.3 Methodology

Cluster analysis is used as the tool for grouping the Delphi statements according to the expertise levels of respondents for this study. The clusters were obtained by pursuing two different strategies; one by applying eigenvector projection technique for attaining a visual representation of the cluster on two-dimensional space, the other one by applying different hierarchical clustering techniques.

The raw data provided by Vision 2023 labels the expertise levels as discrete-ordinal scale data as mentioned in Chapter 5.1.1. It is assumed that the respondents who selected expertise levels 1 and 2 do not have enough knowledge and expertise in the relevant technology; and respondents who selected 3 and 4 are taken into account for the topic, technology field considered. In that sense, data is converted into binary form by assigning the value 0 to expertise levels 1 and 2; the value 1 to expertise levels 3 and 4. If the expertise level is not provided, left empty, it is also labeled as 0.

After obtaining the binary pattern matrix, proximity matrix is obtained. The calculation of proximity matrix requires a selection for a distance or a similarity measure. If an expert gives 0 for two expertise areas it contains no information regarding those two expertise areas, therefore it must be excluded from the calculation. This fact makes the use of Jaccard's coefficient appropriate.

5.3.1 Clusters by Eigenvector Projection

We start by selecting two technology fields (topics) i and j. The total number of respondents that assigned 3 and 4 to the expertise level fields of two topics i and j are calculated and labeled as a. In a similar manner, total number of respondents who assigned 3 or 4 to only one of the topics is counted. The respondents who only have expertise in the area i but not in j are counted as b; and the respondents who only have expertise in the area j but not in i are counted as c. The number of respondents who assigned 0 to the both fields is not counted because such data does not contain any information regarding these two technology fields, topics.

Mathematically the calculation of the Jaccard's coefficient is held with following formula.

$$s(i,j) = a/a+b+c$$

The eigenvectors of the proximity matrix is calculated by using MATLAB 7 tool. The two eigenvectors corresponding to the largest eigenvalues are used for a projection that contains maximum information. However, this projection contains less than 5% of the total information contained in the whole data set. It is found out that the eigenvector projection that contains more than 80% of the total data must use more than 200 dimensions. In other words, a proper representation would require vast amount of dimensions.

To overcome this problem, following Ronde's (2001) work, an eigenvector projection is applied to another proximity matrix which is derived using the previous one. The entries of the new proximity matrix is obtained by

$$d_{ij} = 1/s_{ij}$$

where s_{ij} are the entries of the similarity matrix. In order to overcome the zero division problem, zeroes in the similarity matrix are replaced with a number which is much smaller than the minimum nonzero entry of the similarity matrix. That is to say, we replaced all zero entries with 1/5000 to make sure that maximum distance is represented in the projection as a real number instead of infinity.

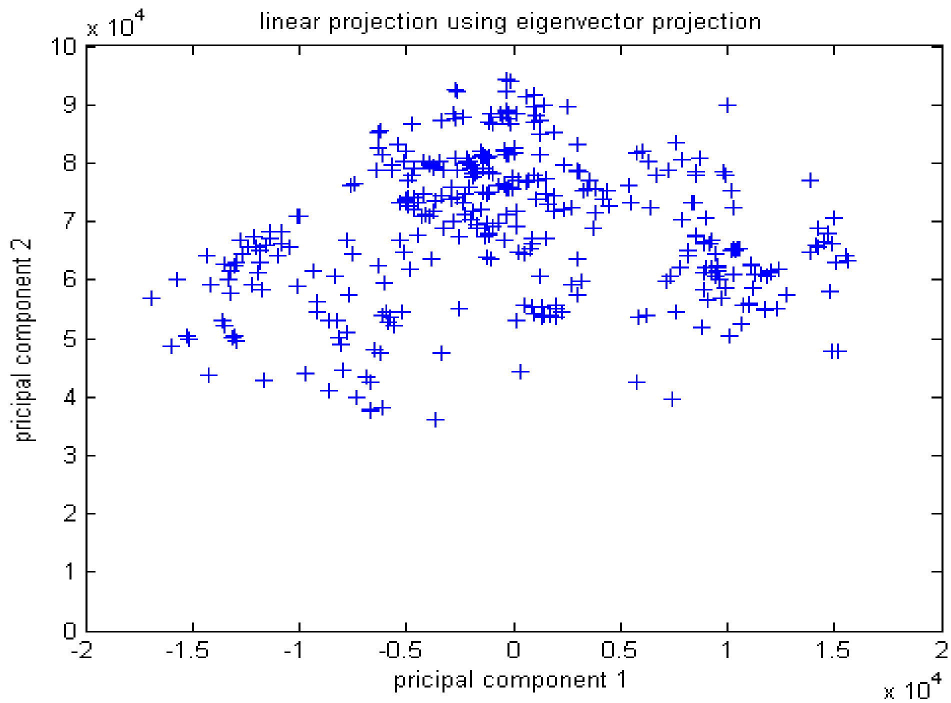

Figure 5.1 Projection in two-dimensional space

The next step is to employ eigenvector projection method explained in Chapter 5.3. The resulting graphic that shows the projected values in 2-dimensional space can be found at Figure 5.1. As in the previous attempt to employ the methodology, too many dimensions have to be used in order to represent the data efficiently. However, analysis is carried out for two dimensions to provide an example. As it can be observed from Figure 5.1, generating clusters by visual examination may not provide reliable results. Therefore, these two dimensional points in the projection plane are clustered by using hierarchical clustering method with squared euclidian distance as the distance metric, however lack of representative power of the resulting clusters makes projection meaningless. The analysis were carried by using MATLAB.

5.3.2 Clusters by Hierarchical Clustering

For hierarchical clustering SPSS 13 was used. In order to guarantee the success of achieving natural clusters hidden in the raw data, different hierarchical techniques were applied to the same data set. Clustering algorithm is fed with binary data and following hierarchical clusters are obtained.

- First version of clusters were acquired by hierarchical cluster with *nearest neighbor (single linkage)*, where similarity measure is *Jaccard's coefficient*.
- Second version of clusters were acquired by hierarchical cluster with *furthest neighbor (complete linkage)*, where similarity measure is *Jaccard's coefficient*.
- Third version of clusters were acquired by hierarchical cluster with *Ward's method*, where distance metric is squared *Euclidian distance*.

Each version is composed of eight sets of clusters; and each set is determined by the number of clusters, from three clusters to ten clusters (a detailed list of topics and their numbers can be found at *http://vizyon2023.tubitak.gov.tr/yayinlar/delfi/*). The clustering process for single linkage and complete linkage methods did not reveal any meaningful clusters. Great majority of topics tended to flock in one cluster, while other few topics were shared among other clusters. A detailed list of clusters by complete linkage and single linkage, and member topics (technologies) can be found in Appendix A1 and Appendix A2. Ward's method on the other hand, clustered topics more significantly. Sequential topics, representing more or less same technologies tended to flock in same clusters. Appendix B1 provides a detailed list of sets of clusters by Ward's method.

For a better understanding of the results of clustering, the sets with four, five and six clusters will be examined here briefly. Most of the topics (technologies) were used in the query of two strategic technologies at the same time; in that sense, two-dimensional tables will be used to provide the knowledge of the distribution of the strategic technologies into clusters. Two-dimensional distribution tables are presented in Appendix B2.

In the first set, there are four clusters. Table 5.5 illustrates the distribution of topics into clusters.

Table 5.5 Four Clusters with Ward's Method

Cluster 1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 283, 289, 291, 302, 344, 345
Cluster 2	25, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 307, 308, 309, 310, 311, 313, 314, 315, 316, 317, 318, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 352
Cluster 3	323, 346, 347, 348, 349, 350, 351, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413
Cluster 4	26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 264, 312, 319, 320, 321, 322

First cluster is an *Information and Communication* technologies dominant cluster; composed of nintynine topics, thirtythree of them take place only in Information and Communication section, and three of them take place only in Defense section. Other than these, Information and Communication field shares eight topics with Defense, three topics with Transportation and two topics with Agriculture fields (Table 5.6).
In general, the first cluster is composed of *Information and Communication*, *Defense Aviation and Space Industry*, *Agriculture and Food*, *Transportation and Tourism*, topics.

Table 5.6 First cluster of Four-cluster set with Ward's Method

		IC	DASI	H&M	T&T	Chem.	C&I	ENR	M&M	ESD	A&F	Tex.	
Cluster 1	IC	**33**	**8**	0	**3**	0	0	0	0	0	**2**	0	
	DASI	0	**3**	0	0	0	0	0	0	0	0	0	
	H&M	0	0	0	0	0	0	0	0	0	0	0	
	T&T	0	0	0	0	0	0	0	0	0	0	0	
	Chem.	0	0	0	0	0	0	0	0	0	0	0	
	C&I	0	0	0	0	0	0	0	0	0	0	0	
	ENR	0	0	0	0	0	0	0	0	0	0	0	
	M&M	0	0	0	0	0	0	0	0	0	0	0	
	ESD	0	0	0	0	0	0	0	0	0	0	0	
	A&F	0	0	0	0	0	0	0	0	0	0	0	
	Tex.	0	0	0	0	0	0	0	0	0	0	0	
		33	**11**	**0**	**3**	**0**	**0**	**0**	**0**	**0**	**2**	**0**	**49**

Second cluster is the largest cluster with twohundred and fiftytwo topics. The cluster is heterogenous in the sense that all technology fields, except Health and Medicine, take place in the cluster with *Machine and Material* technologies slightly dominant. Other than fortyone topics in the cluster which appear only in Machine and Material field; three of the Machine and Material topics also take place in Defense field, six of them also take place in Transportation field, four of them also take place in Chemistry field. Another six of the topics also take place in Energy and Natural Resources field (Table 5.7). In general, the second cluster is composed of

Information and Communication, *Environment and Sustainable Development*, *Energy and Natural Resources*, *Construction and Infrastructure*, *Machine and Material*, *Defense, Aviation and Space Industry*, *Agriculture and Food*, *Transportation and Tourism*, *Textile*, *Chemistry* topics.

Table 5.7 Second cluster of Four-cluster set with Ward's Method

		IC	DASI	H&M	T&T	Chem.	C&I	ENR	M&M	ESD	A&F	Tex.	
Cluster 2	IC	**1**	**1**	0	0	0	0	0	0	0	0	0	
	DASI	0	**27**	0	0	0	0	0	**3**	**1**	0	0	
	H&M	0	0	0	0	**2**	0	0	0	0	0	0	
	T&T	0	0	0	**11**	**1**	**8**	0	**6**	0	0	0	
	Chem.	0	0	0	0	**22**	0	**9**	**4**	0	**1**	**2**	
	C&I	0	0	0	0	0	**15**	**3**	**3**	0	0	0	
	ENR	0	0	0	0	0	0	**23**	**6**	**5**	0	0	
	M&M	0	0	0	0	0	0	0	**41**	0	0	**1**	
	ESD	0	0	0	0	0	0	0	0	**19**	**2**	0	
	A&F	0	0	0	0	0	0	0	0	0	0	0	
	Tex.	0	0	0	0	0	0	0	0	0	0	**35**	
		1	**28**	**0**	**11**	**25**	**23**	**35**	**63**	**25**	**3**	**38**	**252**

Third cluster is mainly *Agriculture and Food* related topics dominated cluster, composed of sixtyeight topics, sixtysix of them take place in Agriculture and Food section. Two of the topics, which are mainly Textile related, are also in Agriculture and Food section (Table 5.8). The cluster is composed of *Agriculture and Food*, and *Textile* topics and all of the topics are related to Agriculture and Food.

Fourth cluster is the smallest cluster with fortyfour topics, mainly dominated by *Health and Medicine* topics. Twentyseven of the topics take place only in Helth and Medicine field, while other seventeen topics again take place in Health and Medicine but also in other fields of Information and Communication, Defense and Machine and Material fields (Table 5.9).

The fourth cluster is composed of *Machine and Material*, *Health and Medicine*, and *Chemistry* topics, and topics are all related to Health and Medicine.

Table 5.8 Third cluster of Four-cluster set with Ward's Method

		IC	DASI	H&M	T&T	Chem.	C&I	ENR	M&M	ESD	A&F	Tex.	
Cluster 3	IC	0	0	0	0	0	0	0	0	0	0	0	
	DASI	0	0	0	0	0	0	0	0	0	0	0	
	H&M	0	0	0	0	0	0	0	0	0	0	0	
	T&T	0	0	0	0	0	0	0	0	0	0	0	
	Chem.	0	0	0	0	0	0	0	0	0	**1**	0	
	C&I	0	0	0	0	0	0	0	0	0	0	0	
	ENR	0	0	0	0	0	0	0	0	0	0	0	
	M&M	0	0	0	0	0	0	0	0	0	0	0	
	ESD	0	0	0	0	0	0	0	0	0	0	0	
	A&F	0	0	0	0	0	0	0	0	0	**65**	**2**	
	Tex.	0	0	0	0	0	0	0	0	0	0	0	
		0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**66**	**2**	**68**

Table 5.9 Fourth cluster of Four-cluster set with Ward's Method

		IC	DASI	H&M	T&T	Chem.	C&I	ENR	M&M	ESD	A&F	Tex.	
Cluster 4	IC	0	0	**7**	0	0	0	0	0	0	0	0	
	DASI	0	0	**1**	0	0	0	0	0	0	0	0	
	H&M	0	0	**27**	0	**8**	0	0	**1**	0	0	0	
	T&T	0	0	0	0	0	0	0	0	0	0	0	
	Chem.	0	0	0	0	0	0	0	0	0	0	0	
	C&I	0	0	0	0	0	0	0	0	0	0	0	
	ENR	0	0	0	0	0	0	0	0	0	0	0	
	M&M	0	0	0	0	0	0	0	0	0	0	0	
	ESD	0	0	0	0	0	0	0	0	0	0	0	
	A&F	0	0	0	0	0	0	0	0	0	0	0	
	Tex.	0	0	0	0	0	0	0	0	0	0	0	
		0	**0**	**35**	**0**	**8**	**0**	**0**	**1**	**0**	**0**	**0**	**44**

In the second set, there are five clusters (Table 5.10). First, third and fourth clusters are the same as set with four clusters. In this second set, the previous second cluster is divided into two to form the fifth cluster. The second cluster was the most heterogenous cluster, combining all fields except Health and Medicine so further division of this cluster into two is a rational clustering. Table 5.11 and 5.12 illustrate the new content of second cluster and fifth cluster.

Table 5.10 Five Clusters with Ward's Method

Cluster 1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 283, 289, 291, 302, 344, 345
Cluster 2	78, 79, 80, 81, 82, 83, 84, 85, 86, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 110, 111, 112, 113, 114, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 225, 258, 260, 262, 263, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 308, 309, 311, 313, 314, 315, 316, 317, 318, 324, 325, 326, 327, 328, 329, 330, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 352
Cluster 3	323, 346, 347, 348, 349, 350, 351, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413
Cluster 4	26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 264, 312, 319, 320, 321, 322
Cluster 5	25, 87, 88, 109, 115, 116, 141, 142, 143, 144, 145, 221, 222, 223, 224, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 259, 261, 265, 266, 267, 268, 307, 310, 331

The new second cluster in five-cluster set carries heterogeneity property forward with the largest number of topics (Table 5.11). The cluster is composed of

Environment and Sustainable Development, *Energy and Natural Resources*, *Construction and Infrastructure*, *Machine and Material*, *Defense*, *Aviation and Space Industry*, *Agriculture and Food*, *Transportation and Tourism*, *Chemistry*, *Textile* topics.

Table 5.11 Second cluster of Five-cluster set with Ward's Method

		IC	DASI	H&M	T&T	Chem.	C&I	ENR	M&M	ESD	A&F	Tex.	
Cluster 2	IC	0	**1**	0	0	0	0	0	0	0	0	0	
	DASI	0	**27**	0	0	0	0	0	**1**	**1**	0	0	
	H&M	0	0	0	0	**2**	0	0	0	0	0	0	
	T&T	0	0	0	**11**	**1**	**8**	0	**1**	0	0	0	
	Chem.	0	0	0	0	**22**	0	**8**	**2**	0	**1**	**2**	
	C&I	0	0	0	0	0	**15**	**3**	**1**	0	0	0	
	ENR	0	0	0	0	0	0	**23**	**2**	**5**	0	0	
	M&M	0	0	0	0	0	0	0	**2**	0	0	**1**	
	ESD	0	0	0	0	0	0	0	0	**19**	**2**	0	
	A&F	0	0	0	0	0	0	0	0	0	0	0	
	Tex.	0	0	0	0	0	0	0	0	0	0	**35**	
		0	**28**	**0**	**11**	**25**	**23**	**34**	**9**	**25**	**3**	**38**	**196**

The fifth cluster is a smaller homogenous cluster, mainly composed of *Machine and Material* topics. Thirtynine of the total fiftysix topics appear only in Machine and Material field. The clustering method separated specific Machine and Material field topics from the rest of the mixture of topics.

The third set of clusters presents six clusters derived from the previous five-cluster set (Table 5.13). Clustering process applies a further division of second cluster, which was the most heterogenous one, into more homogenous clusters.

Table 5.12 Fifth cluster of Five-cluster set with Ward's Method

		IC	DASI	H&M	T&T	Chem.	C&I	ENR	M&M	ESD	A&F	Tex.	
Cluster 5	IC	**1**	0	0	0	0	0	0	0	0	0	0	
	DASI	0	0	0	0	0	0	0	**2**	0	0	0	
	H&M	0	0	0	0	0	0	0	0	0	0	0	
	T&T	0	0	0	0	0	0	0	**5**	0	0	0	
	Chem.	0	0	0	0	0	0	**1**	**2**	0	0	0	
	C&I	0	0	0	0	0	0	0	**2**	0	0	0	
	ENR	0	0	0	0	0	0	0	**4**	0	0	0	
	M&M	0	0	0	0	0	0	0	**39**	0	0	0	
	ESD	0	0	0	0	0	0	0	0	0	0	0	
	A&F	0	0	0	0	0	0	0	0	0	0	0	
	Tex.	0	0	0	0	0	0	0	0	0	0	0	
		1	**0**	**0**	**0**	**0**	**0**	**1**	**54**	**0**	**0**	**0**	**56**

Table 5.13 Six Clusters with Ward's Method

Cluster 1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 283, 289, 291, 302, 344, 345
Cluster 2	78, 79, 80, 81, 82, 83, 84, 85, 86, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 110, 111, 112, 113, 114, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 139, 140, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 225, 258, 260, 262, 263, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 308, 309, 313, 314, 315, 316, 317, 318, 324, 325, 326, 327, 328, 329, 330, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343
Cluster 3	323, 346, 347, 348, 349, 350, 351, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413
Cluster 4	26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 264, 312, 319, 320, 321, 322
Cluster 5	25, 87, 88, 109, 115, 116, 141, 142, 143, 144, 145, 221, 222, 223, 224, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 259, 261, 265, 266, 267, 268, 307, 310, 331
Cluster 6	138, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 311, 352

The second cluster continues to hold topics from all fields except Health and Medicine with the dominancy of Textile and Energy and Natural Resources fields (Table 5.14).

Table 5.14 Second cluster of Six-cluster set with Ward's Method

		IC	DASI	H&M	T&T	Chem.	C&I	ENR	M&M	ESD	A&F	Tex.	
Cluster 2	IC	0	**1**	0	0	0	0	0	0	0	0	0	
	DASI	0	**27**	0	0	0	0	0	**1**	0	0	0	
	H&M	0	0	0	0	**2**	0	0	0	0	0	0	
	T&T	0	0	0	**11**	**1**	**8**	0	**1**	0	0	0	
	Chem.	0	0	0	0	**22**	0	**8**	**2**	0	**1**	**2**	
	C&I	0	0	0	0	0	**15**	**3**	**1**	0	0	0	
	ENR	0	0	0	0	0	0	**23**	**2**	**4**	0	0	
	M&M	0	0	0	0	0	0	0	**2**	0	0	**1**	
	ESD	0	0	0	0	0	0	0	0	0	0	0	
	A&F	0	0	0	0	0	0	0	0	0	0	0	
	Tex.	0	0	0	0	0	0	0	0	0	0	**35**	
		0	**28**	**0**	**11**	**25**	**23**	**34**	**9**	**4**	**1**	**38**	**173**

Table 5.15 Sixth cluster of Six-cluster set with Ward's Method

		IC	DASI	H&M	T&T	Chem.	C&I	ENR	M&M	ESD	A&F	Tex.	
Cluster 6	IC	0	0	0	0	0	0	0	0	0	0	0	
	DASI	0	0	0	0	0	0	0	0	**1**	0	0	
	H&M	0	0	0	0	0	0	0	0	0	0	0	
	T&T	0	0	0	0	0	0	0	0	0	0	0	
	Chem.	0	0	0	0	0	0	0	0	0	0	0	
	C&I	0	0	0	0	0	0	0	0	0	0	0	
	ENR	0	0	0	0	0	0	0	0	**1**	0	0	
	M&M	0	0	0	0	0	0	0	0	0	0	0	
	ESD	0	0	0	0	0	0	0	0	**19**	**2**	0	
	A&F	0	0	0	0	0	0	0	0	0	0	0	
	Tex.	0	0	0	0	0	0	0	0	0	0	0	
		0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**21**	**2**	**0**	**23**

The sixth cluster is mainly composed of *Environment and Sustainable Development* related topics. Nineteen of the total of twentythree topics appear only in the field of Environment and Sustainable Development.

The overall picture of this division of clusters into smaller ones gives a clue about the nature and validity of applied clustering algorithm. Clusters, which hold topics nearly from all technology fields, tend to concentrate on specific fields in further steps. In general, for a better understanding of the content of the clusters in different levels, the technology fields of clusters from three-cluster set to ten-cluster set are given in Appendix B2.

5.4 Scenario Building

There are various ways to construct scenarios and some of them were mentioned with case studies in the works of Ringland (1998) and Fahey and Randall (1998). Different approaches have generated different techniques in scenario building. Here, one of the most common of them will be mentioned; scenario building methodology described by Schwartz and Ogilvy (1998).

Scenario planning begins by identifying and developing intuitive questions to ask about the *focal issue* or *decision* that we want to deal with. In that sense, *identification of our focal issue* is the first step in scenario building. The questions to be asked, can be broad or specific. In the next step, brainstorming begins for

1. Identifying the key factors, which could affect the realization of that decision and
2. Primary *driving forces* at macro level; social dynamics, economic issues, political issues, technological issues

After listing these forces, the *predetermined* ones, the ones that we cannot control, should be eliminated. The final list should be ranked in terms of *importance* and *uncertainty*, two of the most important and uncertain key factors and driving forces should be selected

In the next step, an approach should be selected; either *inductive* or *deductive* in order to identify *scenario logics*. Inductive approach again subcategorizes in itself into *significant events* method and *the official future* method. As far as inductive approach is less structured, both variants require scenario group derive general logical principles either by generating stories around significant future events (significant events method) or by agreeing on an official future and searching for the reasons of any possible deviations from that path (the official future method).

Deductive approach, on the other hand, provides a more structured process. The method basis on construction of a 2X2 scenario matrix, the axis of which is the two most critical uncertainties determined upon group decision (Figure 5.2). The matrix then provides quadrants of uncertainty, which refer to potential futures.

Schwartz and Ogilvy (1998) define a number of advantages of using a matrix in scenario building; assuring qualitatively different scenarios, and key factors' existence in all the scenarios.

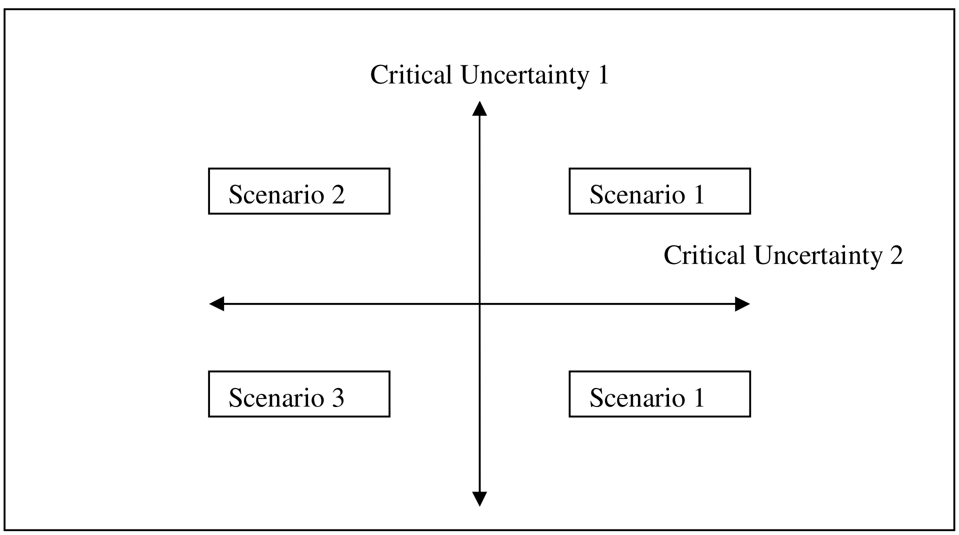

Figure 5.2 Sample Scenario Matrix

Ringland (1998) uses an example of scenario building process, handled by Global Business Network in 1980s to better illustrate the deductive method, an example taken from Art Kleiner's "Why Pacific Gas & Electric Isn't Building Any More Nuclear Power Plant?".

Pacific Gas & Electric was challenged by the question whether to invest in more power plants or pour resources into promoting energy-efficiency. Global Business Network identifies driving forces; environmental movement against nuclear power, economic volatility and immigration. The scenarios revealed that although the company expected people to demand more energy in the future, a very different picture of the world would be an aspect in the future.

As far as this thesis is concerned, performing scenario planning activity as a future looking tool was not aimed. It is important to note that the study, by nature, is not appropriate for developing healthy scenarios in the sense that it lacks the variety of point of views as it did not include any experts, teamwork or a workshop in which scenarios can be developed.

On the other hand, it would be useful to apply before mentioned scenario building method of Schwartz and Ogilvy (1998) to question the future of technology clusters, achieved in the previous section. In order to attain meaningful and manageable clusters, the set with four clusters by Ward's method gave the best result for a better scenario building process.

To sum up the content of the clusters; first cluster is Information and Communication technologies dominant cluster also including Defense, Agriculture and Food, and Transportation and Tourism topics nearly all of them related to Information and Communication field. The second cluster is a crowded cluster, including topics from nearly all fields but exposing a dominancy of Machine and Materials technology topics together with Textile topics. The third cluster is mainly an Agriculture and Food topics dominated cluster; and the fourth cluster is mainly Health and Medicine technologies, including the topics of Chemistry and Machine and Materials topics.

As far as technology clusters identify the nature and content of upcoming innovations, in order to be a leader instead of a follower, to control the change, to develop policies in accordance, to be ready for what the future will hold; it is important to predetermine which of these clusters will dominate the future, which of them will be the major center of innovations and in which of them should the country invest?

In order to obtain a better connection between Delphi and scenario process, the variables according to which the statements were questioned are also included in the scenarios. The inclusion of the variables did not seem to create any unnatural combinations as the variables themselves were selected in line with the needs of the country in Delphi process.

1) Focal issue – Investing in technology clusters.
2) Question – Which technology clusters are likely to lead the future?
3) List of driving forces and key trends –
    - Decisions of corporate networks
    - Policies and budget of government
    - The tax system
    - Number of technical schools
    - Number of universities
    - Price levels
    - EU membership
    - Natural disasters
    - Attitude towards environment
    - Climate change
    - Consumption habits
4) Main uncertainties –
    - We do not know the duration and effects of climate change for sure.
    - We cannot decide government policies, or budget planning.

Our scenario matrix will be like Figure 5.3.

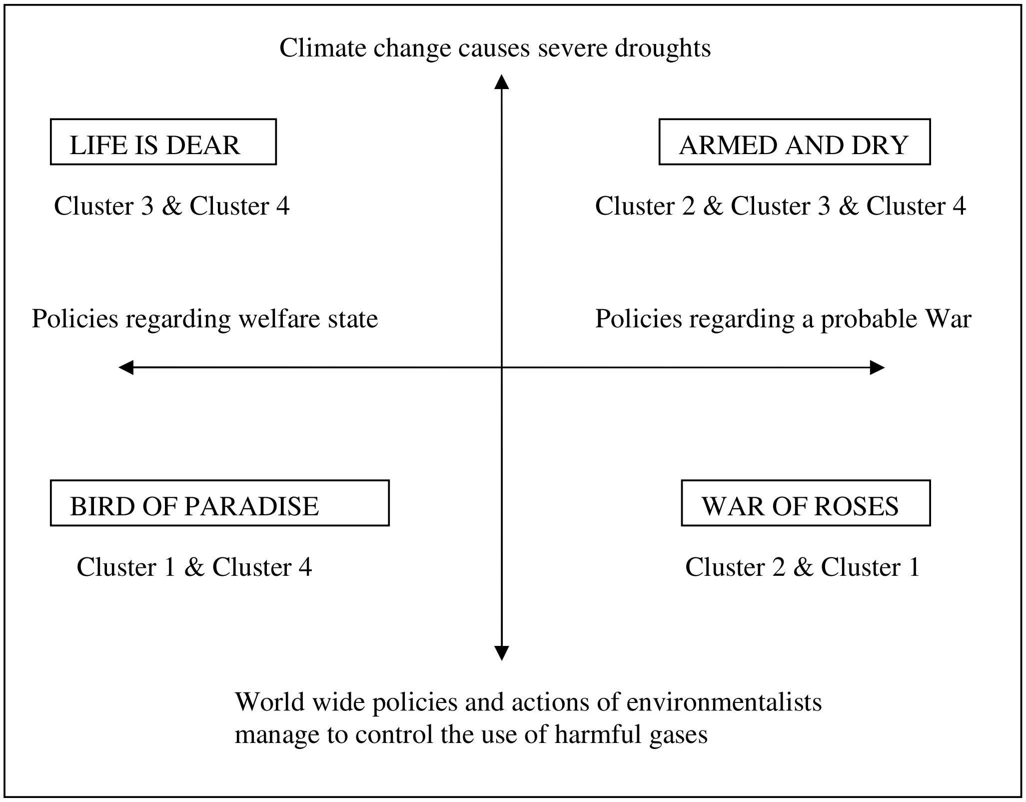

Figure 5.3 Scenario matrix for technologies

<u>Scenario 1 – Armed and Dry</u>

Conditions: Severe droughts, lack of food, new ways of treating new illnesses should be found, panic among people, not only climate change but also “war state” limits the use of resources.

Technologies: Defense technologies will have priority and the related cluster of technologies will support both defense sector and conditions under war and climate change. New types of buildings, new types of clothes against sunrays and new types of military clothings against new generation weapons will be required. Heavy machinery industry will be vital. Besides defense technologies and relevant cluster member technologies, new ways of producing food and pharmaceutics will be supported too. Second and technology clusters will have priority, namely Machine and Materials, Chemistry, Textile, Defense, Agriculture and Food technologies. Fourth technology cluster will have a supporting role, namely Health and Medicine technologies.

Scenario 2 – War of Roses

- Conditions: People are alert. Weather is good but there is war, limiting social aspects of life and leaving people in fear. Everything is about war.
- Technologies: Defense, Machinery and Communication technologies will have priority and the related cluster of technologies will support both defense sector and conditions under war. State has no sympathy for supporting and social needs of people. First and Second clusters will be supported; Information and Communication, Defense, Machine and Materials, Textile, Chemistry technologies.

Scenario 3 – Life is Dear

- Conditions: There is no air, no food, and no land; but there is hope. New kinds of illnesses arose. Governments support welfare policies. Artificial products are required to survive. People have to learn farming on limited lands.
- Technologies: Agriculture and Food technologies have to be developed in order to maintain on limited lands. Artificial vitamins and food products besides new illness treatment systems will be found. Government supports its citizens but costs are high. Third and fourth technology clusters will be supported; namely Agriculture and Food, Health and Medicine and Chemistry technologies.

Scenario 4 – Bird of Paradise

- Conditions: People are happy. Future is hopeful. Government support policies of welfare. Anything is possible and technology serves human for better. Life quality of people is increased. Nothing and everything is luxurious. They want to live in good houses, in green lands, want to travel across countries in minutes and see everything they want without even going there.

- Technologies: Information and Communication technologies will have priority together with Agriculture and Food and Health and Medicine technologies. Especially first and fourth clusters will be supported together with machinery technologies.

CHAPTER 6

CONCLUSION

As far as innovations play important role in determination of economic growth and development, it became vital for nations and organizations to manage the speed and direction of technological changes and innovations. Turkey has failed to catch up with the opportunities resulted from previous technological paradigm shifts. In order to be an active participant and gain a larger share in international markets, Turkey has to develop a better understanding of technology and future. Vision 2023 has provided an initial step for accomplishing a valid and functional policy development and for building an understanding of importance in knowing future by now.

In general, Delphi survey provides a proper environment for interactively collecting the future views of actors of technological systems and systematic use of Delphi results in the building of scenarios provide a more creative tool, presenting various alternatives of future for foresight activities.

Starting from the ideas that innovations occur within clusters and that technologies evolve within systems, the thesis tried to achieve clusters of technologies, which would lead to future innovations in the relevant fields. The data was provided from TUBITAK, Vision 2023 Delphi survey of Turkish Technology Foresight and technologies were clustered in accordance with their common knowledge base. The clusters of technologies then used in the future scenarios of "which technology clusters will be dominant in the future, in which technologies should we invest?" for 2023.

The thesis suggested that common knowledge base would result in the clusters of relevant technologies and the study supported the thesis. Clustering the Delphi statements according to the respondents' expertise levels in relevant fields yielded the result of clusters with common knowledge base. The clusters indicated reasonable and manageable results. Rather than studying technology fields separately, studying them in relation with each other provides a more realistic ground on which the policies can be made. In that sense, the thesis suggests further study on clusters of technologies and investigating the future innovations in relation with the technology clusters rather than questioning the future of technologies one by one.

Scenario writing, on the other hand, provided a more humane methodology for construction of possible futures. Rather than more mechanical presentation of road maps, scenarios provide creativity and inclusion of details in the planning process. They also preserve space for different possibilities of what the future holds. On the other hand, as the rocess required active participation of experts and brainstorming workshops within larger groups; lack of such feature, the thesis could only build a small sample of scenario. The thesis suggests, as a future study, a workshop of scenario building on the findings of Vision 2023 Delphi study.

In general, the technology clusters obtained in the thesis point out the emerging or already present interdiscipliner fields. It would be beneficial to carry out studies in these fields both in terms of academic studies or joint research programs in the sense that the content of clusters reveals the need for cooperation of academic, public and private sector studies. The scenario study which was carried in the thesis also supports the need for joint research and study in the future of technologies, baring possible alternatives of future world in mind.

REFERENCES

Bradfield, R., Wright, G., Burt, G., Cairns, G. and Van Der Heijden, K. (2005). '*The Origins and Evolution of Scenario Techniques in Long Range Business Planning*', Futures, 37, pp. 795-812

Backer, E. (1995). *Computer-Assisted Reasoning in Cluster Analysis*, Prentice Hall, Hertfordshire

Börjeson, L., Höjer, M., Dreborg, K.H., Ekvall, T., Finnveden, G., (2006). '*Scenario Types and Techniques: Towards a User's Guide*'. Futures, 38, pp. 723-739

Campbell, F.J.,(1963). '*The University / Business Research Networks in Science and Technology Knowledge Production Trends in the US, EU and Japan'* in Knowledge Creation Diffusion and Use in Innovation Networks and Knowledge Clusters, pp 67-100

Carlsson, B. and Eliasson, G., (2001). '*Industrial Dynamics and Endogenous Growth*', DRUID

Coates, J.P., (2000). '*From My Perspective: Scenario Planning*', Technological Forecasting and Social Change, 65, pp. 115-123

Cuhls, K. (2003). '*From Forecasting to Foresight Processes – New Participative Foresight Acrivities in Germany*', Journal of Forecasting, 22, pp. 93-111

Dalkey, N.C., (1967). '*Delphi*', RAND, (P-3704)

Dalkey, N.C., (1969). '*The Delphi Method: An Experimental Study of Group Opinion*', RAND, (RM-5888-PR)

Durand, T. (2003). '*Twelve Lessons from 'Key Technologies 2005': the French Tecchnology Foresight Exercise*', Journal of Forecasting, 22, pp. 161-177

Everitt, B., (1974). Cluster Analysis, Heinemann Educational Books, London

Fagerberg, J., (2003). '*Schumpeter and Revival of Evolutionary Economics: An Appraisal of the Literature'*, Journal of Evolutionary Economics, 13, pp.125-159

Fahey, L. and Randall, R. M., (1998). '*What is Scenario Learning?*' in Fahey, L. & Randall, R. M., (Eds.) Learning From The Future: Competitive Foresight Scenarios, John Wiley & Sons, New York

FOREN Network (2001). *A Practical Guide to Regional Foresight*, European Commission – STRATA Programme

Freeman, C.; Clark, C. and Soete, L., (1982). *Unemployment and Technical Innovation: A Study of Long Waves in Economic Development*, Frances Pinter, London

Godet, M., (1995) '*Global Scenarios: Morphological and Probabillity Analysis*' in European Commission Joint Reseaarch Center, Scenario Building: Convergences and Differences, Workshop Proceedings, Spain, EUR-17298-EN

Göker, H.A. (1999). *'Yeni Teknolojiler Sanayide Değişim ve Yeni Sanayiler'*, MMO Sanayi Kongresi

Grupp, H., Linstone, H.A. (1999). '*National technology Foresight Activities Around the Globe: Resurrection and New Paradigms*', Technological Forecasting and Social Change, 60, pp. 85-94

Jain, A.K. and Dubes, R.C. (1988). *Algorithms for Clustering Data*, Prentice Hall, Englewood Cliffs

Janowitz, M.F. (2002). *Short Course: A Combinational Introduction to Cluster Analysis*, CSNA

Johnson, R.A. and Wichern, D.W.W. (1998). *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey

Kaufman, L. and Rousseeuw, P.J. (2005). *Finding Groups in Data: An Introduction to Cluster Analysis*, John Wiley & Sons, New Jersey

Linstone, H.A. and Turoff, M., (2002). *The Delphi Method: Techniques and Applications*, http://is.njit.edu/pubs/delphibook/

de Loe, R., (1995). '*Exploring Complex Policy Questions Using the Policy Delphi: A Multi-round, Interactive Survey Method*', Applied Geography, 15 (1), pp.53-68

Malaska, P., (1995). '*Surveys of the Use of the Multiple Scenario Approach in Big European Companies since 1973*' in Scenario Building: Convergences and Differences, Profutures Workshop, Workshop proceedings

Martin, B. R. (2002). '*Technology Foresight in a Rapidly Globalizing Economy*', SPRU, Brighton

Martin, B.R., Johnston, R. (1999). '*Technology Foresight for Wiring Up the National Innovation System: Experiences in Britain, Australia, and New Zealand*', Technological Forecasting and Social Change, 60, pp.37-54

Masini, E.B. and Vasquez, J.M., (2000). '*Scenarios as Seen from a Human and Social Perspective*', Technological Forecasting and Social Change, 65, pp.49-66

McMaster, M. (1996). '*Foresight: Exploring the Structure of the Future*', Long Range Planning, 29, pp.149-155

Nelson, R.R. and Winter, S.G., (2002). '*Evolutionary Theorizing in Economics'*, Journal of Economic Perspectives, 16 (2), pp. 23-46

Perez, C.,(1983). '*Structural Change and Assimilation of New Technologies in the Economic and Social Systems"*, Futures, 15, pp.357-375

Perez, C., (1986). '*The New Technologies: An Integrated View*' http://www.carlotaperez.org/papers/basic-anintegratedview.htm

Perez, C., (2004). '*Technological Revolutions, Paradigm Shifts and Socio-Institutional Change*' http://www.carlotaperez.org/papers/basic-technologicalrevolutionsparadigm.htm

Perrottet, C.M., (1998). '*Testing Your Strategies in Scenarios*' in Fahey, L. & Randall, R. M., (Eds.) Learning From The Future: Competitive Foresight Scenarios, John Wiley & Sons, New York

Pill, J., (1971). '*The Delphi Method: Substance, Context, A Critique and An Annotated Bibliography*', Socio-Econ. Plan. Sci., 5, pp. 57-71

Ringland, G. (1998). *Scenario Planning: Managing for the Future*, John Wiley & Sons, Chichester

Ronde, P., (2001). '*Technological Clusters with a Knowledge-based Principle: Evidence from a Delphi Investigation in the French Case of the Life Sciences*', Research Policy, 30, pp.1041-1057

Stead, D. and Banister, D., (2003). '*Transport Policy Scenario-Building*', Transportation Planning and Technology, 26 (6), pp. 513-536

Schwartz, P. and Ogilvy, J.A., (1995). '*Plotting Your Scenarios*' in Fahey, L. & Randall, R. M., (Eds.) Learning From The Future: Competitive Foresight Scenarios, John Wiley & Sons, New York

Schumpeter, J.A., (1939). *Business Cycles; a Theoretical, Historical, and Statistical Analysis of the Capitalist System*, McGraw-Hill, New York, London

Thomas, C.W., (1998). '*Scenario-Based Planning for Technology Investments*' in Fahey, L. & Randall, R. M., (Eds.) Learning From The Future: Competitive Foresight Scenarios, John Wiley & Sons, New York

TUBITAK, (2004). *Ulusal Bilim ve Teknoloji Politikaları 2003-2023 Strateji Belgesi*, Versiyon 19

TUBITAK, http://vizyon2023.tubitak.gov.tr/yayinlar/delfi/ , 23.07.2004

Tümer, T. (2002). '*Technology Foresight in Turkey*', PowerPoint presentation in Eforesee Workshop on Foresight Basics, Malta

Webster , A. (1999). '*Technologies in Transition, Policies in Transition: Foresight in the Risk Society*', Technovation, 19, pp. 413-421

Wilson, I., (1998). '*Mental Maps of the Future: An Intuitive Logics Approach to Scenarios* ' in Fahey, L. & Randall, R. M., (Eds.) Learning From The Future: Competitive Foresight Scenarios, John Wiley & Sons, New York

Von Schomberg, R., Pereira, A. G. and Funtowicz, S. (2005). *Deliberating Foresight Knowledge for Policy and Foresight Knowledge Assessment: A Working Document from the European Commission Services*

APPENDICES

APPENDIX A1

Hierarchical Clustering with Single linkage; 10 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 331, 330, 329, 328, 327, 326, 325, 324, 323, 322, 321, 320, 319, 318, 317, 316, 315, 314, 313, 312, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 264, 263, 262, 261, 259, 257, 256, 255, 254, 253, 252, 251, 250, 249, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 238, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 227, 226, 224, 223, 222, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 163, 162, 161, 160, 159, 158, 157, 156, 155, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 144, 143, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 116, 115, 114, 113, 112, 111, 110, 109, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 45

Cluster 3: 86

Cluster 4: 97

Cluster 5: 225

Cluster 6: 258

Cluster 7: 260

Cluster 8: 301

Cluster 9: 345, 344

Cluster 10: 398, 397, 396, 395

Hierarchical Clustering with Single linkage; 9 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 331, 330, 329, 328, 327, 326, 325, 324, 323, 322, 321, 320, 319, 318, 317, 316, 315, 314, 313, 312, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 264, 263, 262, 261, 259, 257, 256, 255, 254, 253, 252, 251, 250, 249, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 238, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 227, 226, 224, 223, 222, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 163, 162, 161, 160, 159, 158, 157, 156, 155, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 144, 143, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 116, 115, 114, 113, 112, 111, 110, 109, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 45

Cluster 3: 86

Cluster 4: 97

Cluster 5: 225

Cluster 6: 258

Cluster 7: 260

Cluster 8: 301

Cluster 9: 398, 397, 396, 395

Hierarchical Clustering with Single linkage; 8 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 331, 330, 329, 328, 327, 326, 325, 324, 323, 322, 321, 320, 319, 318, 317, 316, 315, 314, 313, 312, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 264, 263, 262, 261, 259, 257, 256, 255, 254, 253, 252, 251, 250, 249, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 238, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 227, 226, 224, 223, 222, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 163, 162, 161, 160, 159, 158, 157, 156, 155, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 144, 143, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 116, 115, 114, 113, 112, 111, 110, 109, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 45

Cluster 3: 97

Cluster 4: 225

Cluster 5: 258

Cluster 6: 260

Cluster 7: 301

Cluster 8: 398, 397, 396, 395

Hierarchical Clustering with Single linkage; 7 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 331, 330, 329, 328, 327, 326, 325, 324, 323, 322, 321, 320, 319, 318, 317, 316, 315, 314, 313, 312, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 264, 263, 262, 261, 259, 258, 257, 256, 255, 254, 253, 252, 251, 250, 249, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 238, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 227, 226, 224, 223, 222, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 163, 162, 161, 160, 159, 158, 157, 156, 155, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 144, 143, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 116, 115, 114, 113, 112, 111, 110, 109, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 45

Cluster 3: 97

Cluster 4: 225

Cluster 5: 260

Cluster 6: 301

Cluster 7: 398, 397, 396, 395

Hierarchical Clustering with Single linkage; 6 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 331, 330, 329, 328, 327, 326, 325, 324, 323, 322, 321, 320, 319, 318, 317, 316, 315, 314, 313, 312, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 264, 263, 262, 261, 259, 258, 257, 256, 255, 254, 253, 252, 251, 250, 249, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 238, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 227, 226, 224, 223, 222, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 163, 162, 161, 160, 159, 158, 157, 156, 155, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 144, 143, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 116, 115, 114, 113, 112, 111, 110, 109, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 97

Cluster 3: 225

Cluster 4: 260

Cluster 5: 301

Cluster 6: 398, 397, 396, 395

Hierarchical Clustering with Single linkage; 5 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 331, 330, 329, 328, 327, 326, 325, 324, 323, 322, 321, 320, 319, 318, 317, 316, 315, 314, 313, 312, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 264, 263, 262, 261, 259, 258, 257, 256, 255, 254, 253, 252, 251, 250, 249, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 238, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 227, 226, 224, 223, 222, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 163, 162, 161, 160, 159, 158, 157, 156, 155, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 144, 143, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 116, 115, 114, 113, 112, 111, 110, 109, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 97

Cluster 3: 225

Cluster 4: 260

Cluster 5: 301

Hierarchical Clustering with Single linkage; 4 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 331, 330, 329, 328, 327, 326, 325, 324, 323, 322, 321, 320, 319, 318, 317, 316, 315, 314, 313, 312, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 301, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 264, 263, 262, 261, 259, 258, 257, 256, 255, 254, 253, 252, 251, 250, 249, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 238, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 227, 226, 224, 223, 222, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 163, 162, 161, 160, 159, 158, 157, 156, 155, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 144, 143, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 116, 115, 114, 113, 112, 111, 110, 109, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 97

Cluster 3: 225

Cluster 4: 260

Hierarchical Clustering with Single linkage; 3 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 331, 330, 329, 328, 327, 326, 325, 324, 323, 322, 321, 320, 319, 318, 317, 316, 315, 314, 313, 312, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 301, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 264, 263, 262, 261, 259, 258, 257, 256, 255, 254, 253, 252, 251, 250, 249, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 238, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 227, 226, 225, 224, 223, 222, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 163, 162, 161, 160, 159, 158, 157, 156, 155, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 144, 143, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 116, 115, 114, 113, 112, 111, 110, 109, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 97

Cluster 3: 260

APPENDIX A2

Hierarchical Clustering with Complete linkage; 10 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 351, 350, 349, 348, 347, 346, 345, 344, 323, 316, 309, 308, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 260, 258, 257, 255, 238, 227, 224, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 169, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 331, 254, 253, 252, 251, 250, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 223, 222, 116, 115, 109, 25

Cluster 3: 330, 322, 321, 320, 319, 312, 301, 264, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 26

Cluster 4: 96, 95, 92, 91, 90, 89, 85, 84, 83, 81, 80, 79, 78

Cluster 5: 163, 162, 161, 160, 159, 158, 157, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 86, 82

Cluster 6: 310, 307, 268, 267, 266, 265, 261, 259, 256, 249, 226, 225, 144, 143, 88, 87

Cluster 7: 118, 117, 114, 113, 112, 111, 110, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 94, 93

Cluster 8: 97

Cluster 9: 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 329, 328, 327, 326, 325, 324, 318, 317, 315, 314, 313, 263, 262, 156, 155

Cluster 10: 352, 311, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 168, 167, 166, 165, 164

Hierarchical Clustering with Complete linkage; 9 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 323, 316, 311, 309, 308, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 260, 258, 257, 255, 238, 227, 224, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 331, 254, 253, 252, 251, 250, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 223, 222, 116, 115, 109, 25

Cluster 3: 330, 322, 321, 320, 319, 312, 301, 264, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 26

Cluster 4: 96, 95, 92, 91, 90, 89, 85, 84, 83, 81, 80, 79, 78

Cluster 5: 163, 162, 161, 160, 159, 158, 157, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 86, 82

Cluster 6: 310, 307, 268, 267, 266, 265, 261, 259, 256, 249, 226, 225, 144, 143, 88, 87

Cluster 7: 118, 117, 114, 113, 112, 111, 110, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 94, 93

Cluster 8: 97

Cluster 9: 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 329, 328, 327, 326, 325, 324, 318, 317, 315, 314, 313, 263, 262, 156, 155

Hierarchical Clustering with Complete linkage; 8 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 329, 328, 327, 326, 325, 324, 323, 318, 317, 316, 315, 314, 313, 311, 309, 308, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 263, 262, 260, 258, 257, 255, 238, 227, 224, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 156, 155, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 331, 254, 253, 252, 251, 250, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 223, 222, 116, 115, 109, 25

Cluster 3: 330, 322, 321, 320, 319, 312, 301, 264, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 26

Cluster 4: 96, 95, 92, 91, 90, 89, 85, 84, 83, 81, 80, 79, 78

Cluster 5: 163, 162, 161, 160, 159, 158, 157, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 86, 82

Cluster 6: 310, 307, 268, 267, 266, 265, 261, 259, 256, 249, 226, 225, 144, 143, 88, 87

Cluster 7: 118, 117, 114, 113, 112, 111, 110, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 94, 93

Cluster 8: 97

Hierarchical Clustering with Complete linkage; 7 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 329, 328, 327, 326, 325, 324, 323, 318, 317, 316, 315, 314, 313, 311, 309, 308, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 263, 262, 260, 258, 257, 255, 238, 227, 224, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 156, 155, 97, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 331, 254, 253, 252, 251, 250, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 223, 222, 116, 115, 109, 25

Cluster 3: 330, 322, 321, 320, 319, 312, 301, 264, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 26

Cluster 4: 96, 95, 92, 91, 90, 89, 85, 84, 83, 81, 80, 79, 78

Cluster 5: 163, 162, 161, 160, 159, 158, 157, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 86, 82

Cluster 6: 310, 307, 268, 267, 266, 265, 261, 259, 256, 249, 226, 225, 144, 143, 88, 87

Cluster 7: 118, 117, 114, 113, 112, 111, 110, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 94, 93

Hierarchical Clustering with Complete linkage; 6 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 329, 328, 327, 326, 325, 324, 323, 318, 317, 316, 315, 314, 313, 311, 309, 308, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 263, 262, 260, 258, 257, 255, 238, 227, 224, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 156, 155, 118, 117, 114, 113, 112, 111, 110, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 97, 94, 93, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1

Cluster 2: 331, 254, 253, 252, 251, 250, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 223, 222, 116, 115, 109, 25

Cluster 3: 330, 322, 321, 320, 319, 312, 301, 264, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 26

Cluster 4: 96, 95, 92, 91, 90, 89, 85, 84, 83, 81, 80, 79, 78

Cluster 5: 163, 162, 161, 160, 159, 158, 157, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 86, 82

Cluster 6: 310, 307, 268, 267, 266, 265, 261, 259, 256, 249, 226, 225, 144, 143, 88, 87

Hierarchical Clustering with Complete linkage; 5 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 329, 328, 327, 326, 325, 324, 323, 318, 317, 316, 315, 314, 313, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 263, 262, 261, 260, 259, 258, 257, 256, 255, 249, 238, 227, 226, 225, 224, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 156, 155, 144, 143, 118, 117, 114, 113, 112, 111, 110, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 97, 94, 93, 88, 87, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1

Cluster 2: 331, 254, 253, 252, 251, 250, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 223, 222, 116, 115, 109, 25

Cluster 3: 330, 322, 321, 320, 319, 312, 301, 264, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 26

Cluster 4: 96, 95, 92, 91, 90, 89, 85, 84, 83, 81, 80, 79, 78

Cluster 5: 163, 162, 161, 160, 159, 158, 157, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 86, 82

Hierarchical Clustering with Complete linkage; 4 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 329, 328, 327, 326, 325, 324, 323, 318, 317, 316, 315, 314, 313, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 263, 262, 261, 260, 259, 258, 257, 256, 255, 249, 238, 227, 226, 225, 224, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 163, 162, 161, 160, 159, 158, 157, 156, 155, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 144, 143, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 114, 113, 112, 111, 110, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 97, 94, 93, 88, 87, 86, 82, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 331, 254, 253, 252, 251, 250, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 223, 222, 116, 115, 109, 25

Cluster 3: 330, 322, 321, 320, 319, 312, 301, 264, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 26

Cluster 4: 96, 95, 92, 91, 90, 89, 85, 84, 83, 81, 80, 79, 78

Hierarchical Clustering with Complete linkage; 3 Clusters

Cluster 1: 413, 412, 411, 410, 409, 408, 407, 406, 405, 404, 403, 402, 401, 400, 399, 398, 397, 396, 395, 394, 393, 392, 391, 390, 389, 388, 387, 386, 385, 384, 383, 382, 381, 380, 379, 378, 377, 376, 375, 374, 373, 372, 371, 370, 369, 368, 367, 366, 365, 364, 363, 362, 361, 360, 359, 358, 357, 356, 355, 354, 353, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 332, 329, 328, 327, 326, 325, 324, 323, 318, 317, 316, 315, 314, 313, 311, 310, 309, 308, 307, 306, 305, 304, 303, 302, 300, 299, 298, 297, 296, 295, 294, 293, 292, 291, 290, 289, 288, 287, 286, 285, 284, 283, 282, 281, 280, 279, 278, 277, 276, 275, 274, 273, 272, 271, 270, 269, 268, 267, 266, 265, 263, 262, 261, 260, 259, 258, 257, 256, 255, 249, 238, 227, 226, 225, 224, 221, 220, 219, 218, 217, 216, 215, 214, 213, 212, 211, 210, 209, 208, 207, 206, 205, 204, 203, 202, 201, 200, 199, 198, 197, 196, 195, 194, 193, 192, 191, 190, 189, 188, 187, 186, 185, 184, 183, 182, 181, 180, 179, 178, 177, 176, 175, 174, 173, 172, 171, 170, 169, 168, 167, 166, 165, 164, 163, 162, 161, 160, 159, 158, 157, 156, 155, 154, 153, 152, 151, 150, 149, 148, 147, 146, 145, 144, 143, 142, 141, 140, 139, 138, 137, 136, 135, 134, 133, 132, 131, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 114, 113, 112, 111, 110, 108, 107, 106, 105, 104, 103, 102, 101, 100, 99, 98, 97, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1,

Cluster 2: 331, 254, 253, 252, 251, 250, 248, 247, 246, 245, 244, 243, 242, 241, 240, 239, 237, 236, 235, 234, 233, 232, 231, 230, 229, 228, 223, 222, 116, 115, 109, 25

Cluster 3: 330, 322, 321, 320, 319, 312, 301, 264, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 26

APPENDIX B1

Hierarchical Clustering with Ward's Method; 10 Clusters

Cluster1: 1, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 289, 291, 302, 344, 345

Cluster 2: 78, 79, 80, 81, 83, 84, 85, 86, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 110, 111, 112, 113, 114, 117, 118, 133, 134, 135, 136, 137, 146, 147, 148, 149, 150, 151, 152, 153, 154, 163, 185, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 225, 258, 260, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 308, 309,

Cluster 3: 323, 348, 349, 350, 351, 353, 354, 359, 360, 364, 366, 368, 369, 370, 371, 372, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413

Cluster 4: 26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 264, 312, 319, 320, 321, 322

Cluster 5: 25, 87, 88, 109, 115, 116, 141, 142, 143, 144, 145, 221, 222, 223, 224, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 259, 261, 265, 266, 267, 268, 307, 310, 331

Cluster 6: 138, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 311, 352

Cluster 7: 124, 125, 126, 127, 128, 129, 139, 140, 155, 156, 184, 186, 262, 263, 313, 314, 315, 316, 317, 318, 324, 325, 326, 327, 328, 329, 330, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343

Cluster 8: 346, 347, 355, 356, 357, 358, 361, 362, 363, 365, 367, 373

Cluster 9: 82, 119, 120, 121, 122, 123, 130, 131, 132, 157, 158, 159, 160, 161, 162

Cluster 10: 2, 3, 4, 283

Hierarchical Clustering with Ward's Method; 9 Clusters

Cluster 1: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 283, 289, 291, 302, 344, 345

Cluster 2: 78, 79, 80, 81, 83, 84, 85, 86, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 110, 111, 112, 113, 114, 117, 118, 133, 134, 135, 136, 137, 146, 147, 148, 149, 150, 151, 152, 153, 154, 163, 185, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 225, 258, 260, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 308, 309

Cluster 3: 323, 348, 349, 350, 351, 353, 354, 359, 360, 364, 366, 368, 369, 370, 371, 372, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413

Cluster 4: 26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 264, 312, 319, 320, 321, 322

Cluster 5: 25, 87, 88, 109, 115, 116, 141, 142, 143, 144, 145, 221, 222, 223, 224, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 259, 261, 265, 266, 267, 268, 307, 310, 331

Cluster 6: 138, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 311, 352

Cluster 7: 124, 125, 126, 127, 128, 129, 139, 140, 155, 156, 184, 186, 262, 263, 313, 314, 315, 316, 317, 318, 324, 325, 326, 327, 328, 329, 330, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343

Cluster 8: 346, 347, 355, 356, 357, 358, 361, 362, 363, 365, 367, 373

Cluster 9: 82, 119, 120, 121, 122, 123, 130, 131, 132, 157, 158, 159, 160, 161, 162

Hierarchical Clustering with Ward's Method; 8 Clusters:

Cluster 1: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 283, 289, 291, 302, 344, 345

Cluster 2: 25, 87, 88, 109, 115, 116, 141, 142, 143, 144, 145, 221, 222, 223, 224, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 259, 261, 265, 266, 267, 268, 307, 310, 331

Cluster 3: 323, 348, 349, 350, 351, 353, 354, 359, 360, 364, 366, 368, 369, 370, 371, 372, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413

Cluster 4: 26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 264, 312, 319, 320, 321, 322

Cluster 5: 78, 79, 80, 81, 83, 84, 85, 86, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 110, 111, 112, 113, 114, 117, 118, 133, 134, 135, 136, 137, 146, 147, 148, 149, 150, 151, 152, 153, 154, 163, 185, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 225, 258, 260, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 308, 309,

Cluster 6: 138, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 311, 352

Cluster 7: 82, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 139, 140, 155, 156, 157, 158, 159, 160, 161, 162, 184, 186, 262, 263, 313, 314, 315, 316, 317, 318, 324, 325, 326, 327, 328, 329, 330, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343

Cluster 8: 346, 347, 355, 356, 357, 358, 361, 362, 363, 365, 367, 373

Hierarchical Clustering with Ward's Method; 7 Clusters

Cluster 1: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 283, 289, 291, 302, 344, 345

Cluster 2: 78, 79, 80, 81, 83, 84, 85, 86, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 110, 111, 112, 113, 114, 117, 118, 133, 134, 135, 136, 137, 146, 147, 148, 149, 150, 151, 152, 153, 154, 163, 185, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 225, 258, 260, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 308, 309,

Cluster 3: 323, 346, 347, 348, 349, 350, 351, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413,

Cluster 4: 26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 264, 312, 319, 320, 321, 322

Cluster 5: 25, 87, 88, 109, 115, 116, 141, 142, 143, 144, 145, 221, 222, 223, 224, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 259, 261, 265, 266, 267, 268, 307, 310, 331

Cluster 6: 138, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 311, 352

Cluster 7: 82, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 139, 140, 155, 156, 157, 158, 159, 160, 161, 162, 184, 186, 262, 263, 313, 314, 315, 316, 317, 318, 324, 325, 326, 327, 328, 329, 330, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343

Hierarchical Clustering with Ward's Method; 6 Clusters:

Cluster 1: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 283, 289, 291, 302, 344, 345

Cluster 2: 78, 79, 80, 81, 82, 83, 84, 85, 86, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 110, 111, 112, 113, 114, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 139, 140, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 225, 258, 260, 262, 263, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 308, 309, 313, 314, 315, 316, 317, 318, 324, 325, 326, 327, 328, 329, 330, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343

Cluster 3: 323, 346, 347, 348, 349, 350, 351, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413

Cluster 4: 26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 264, 312, 319, 320, 321, 322

Cluster 5: 25, 87, 88, 109, 115, 116, 141, 142, 143, 144, 145, 221, 222, 223, 224, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 259, 261, 265, 266, 267, 268, 307, 310, 331

Cluster 6: 138, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 311, 352

Hierarchical Clustering with Ward's Method; 5 Clusters:

Cluster 1: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 283, 289, 291, 302, 344, 345

Cluster 2: 78, 79, 80, 81, 82, 83, 84, 85, 86, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 110, 111, 112, 113, 114, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 225, 258, 260, 262, 263, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 308, 309, 311, 313, 314, 315, 316, 317, 318, 324, 325, 326, 327, 328, 329, 330, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 352

Cluster 3: 323, 346, 347, 348, 349, 350, 351, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413

Cluster 4: 26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 264, 312, 319, 320, 321, 322

Cluster 5: 25, 87, 88, 109, 115, 116, 141, 142, 143, 144, 145, 221, 222, 223, 224, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 259, 261, 265, 266, 267, 268, 307, 310, 331

Hierarchical Clustering with Ward's Method; 4 Clusters:

Cluster 1: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 283, 289, 291, 302, 344, 345

Cluster 2: 25, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 307, 308, 309, 310, 311, 313, 314, 315, 316, 317, 318, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 352

Cluster 3: 323, 346, 347, 348, 349, 350, 351, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413

Cluster 4: 26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 264, 312, 319, 320, 321, 322

Hierarchical Clustering with Ward's Method; 3 Clusters:

Cluster 1: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 275, 276, 277, 279, 282, 283, 289, 291, 302, 344, 345

Cluster 2: 25, 26, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 278, 280, 281, 284, 285, 286, 287, 288, 290, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 352

Cluster 3: 323, 346, 347, 348, 349, 350, 351, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413,

APPENDIX B2

Distribution of technology fields in ten-cluster set

Cluster 1; Information and Communication Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics. The dominant field is *Information and Communication*.

Cluster 2; Environment and Sustainable Development Topics, Energy and Natural Resources Topics, Construction and Infrastructure Topics, Machine and Material Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Chemistry Topics, Textile Topics. The dominant field is *Textile*.

Cluster 3; Agriculture and Food Topics, Textile Topics. The dominant field is *Agriculture and Food*

Cluster 4; Machine and Material Topics, Health and Medicine Topics, Chemistry Topics. The dominant field is *Health and Medicine*.

Cluster 5; Information and Communication Topics, Energy and Natural Resources Topics, Machine and Material Topics. The dominant field is *Machine and Materials*.

Cluster 6; Environment and Sustainable Development Topics, Agriculture and Food Topics. The dominant field is *Environment and Sustainable Development*.

Cluster 7; Energy and Natural Resources Topics, Machine and Material Topics, Agriculture and Food Topics, Chemistry Topics, Textile Topics. The dominant field is *Chemistry*.

Cluster 8; *Agriculture and Food Topics*.

Cluster 9: Environment and Sustainable Development Topics, Energy and Natural Resources Topics, Construction and Infrastructure Topics. The dominant field is *Energy and Natural Resources*.

Cluster 10: Information and Communication Topics, Defense, Aviation and Space Industry Topics. The dominant field is *Information and Communication*.

Distribution of technology fields in nine-cluster set

Cluster 1; Information and Communication Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Agriculture and Food Topics. The dominant field is *Information and Communication*.

Cluster 2; Environment and Sustainable Development Topics, Energy and Natural Resources Topics, Construction and Infrastructure Topics, Machine and Material Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Chemistry Topics, Textile Topics. The dominant field is *Textile*.

Cluster 3; Agriculture and Food Topics, Textile Topics. The dominant field is *Agriculture and Food*

Cluster 4; Machine and Material Topics, Health and Medicine Topics, Chemistry Topics. The dominant field is *Health and Medicine*.

Cluster 5; Information and Communication Topics, Energy and Natural Resources Topics, Machine and Material Topics. The dominant field is *Machine and Materials*.

Cluster 6; Environment and Sustainable Development Topics, Agriculture and Food Topics. The dominant field is *Environment and Sustainable Development*.

Cluster 7; Energy and Natural Resources Topics, Machine and Material Topics, Agriculture and Food Topics, Chemistry Topics, Textile Topics. The dominant field is *Chemistry*.

Cluster 8; *Agriculture and Food Topics*.

Cluster 9: Environment and Sustainable Development Topics, Energy and Natural Resources Topics, Construction and Infrastructure Topics. The dominant field is *Energy and Natural Resources*.

Distribution of technology fields in eight-cluster set

Cluster 1; Information and Communication Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Agriculture and Food Topics. The dominant field is *Information and Communication*.

Cluster 2; Environment and Sustainable Development Topics, Energy and Natural Resources Topics, Construction and Infrastructure Topics, Machine and Material Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Chemistry Topics, Textile Topics. The dominant field is *Textile*.

Cluster 3; Agriculture and Food Topics, Textile Topics. The dominant field is *Agriculture and Food*

Cluster 4; Machine and Material Topics, Health and Medicine Topics, Chemistry Topics. The dominant field is *Health and Medicine*.

Cluster 5; Information and Communication Topics, Energy and Natural Resources Topics, Machine and Material Topics. The dominant field is *Machine and Materials*.

Cluster 6; Environment and Sustainable Development Topics, Agriculture and Food Topics. The dominant field is *Environment and Sustainable Development*.

Cluster 7; Energy and Natural Resources Topics, Machine and Material Topics, Agriculture and Food Topics, Chemistry Topics, Textile Topics, Environment and Sustainable Development Topics, Construction and Infrastructure Topics. The dominant field is *Chemistry*.

Cluster 8; *Agriculture and Food Topics*.

Distribution of technology fields in seven-cluster set

Cluster 1; Information and Communication Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Agriculture and Food Topics. The dominant field is *Information and Communication*.

Cluster 2; Environment and Sustainable Development Topics, Energy and Natural Resources Topics, Construction and Infrastructure Topics, Machine and Material Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Chemistry Topics, Textile Topics. The dominant field is *Textile*.

Cluster 3; Agriculture and Food Topics, Textile Topics. The dominant field is *Agriculture and Food*

Cluster 4; Machine and Material Topics, Health and Medicine Topics, Chemistry Topics. The dominant field is *Health and Medicine*.

Cluster 5; Information and Communication Topics, Energy and Natural Resources Topics, Machine and Material Topics. The dominant field is *Machine and Materials*.

Cluster 6; Environment and Sustainable Development Topics, Agriculture and Food Topics. The dominant field is *Environment and Sustainable Development*.

Cluster 7; Energy and Natural Resources Topics, Machine and Material Topics, Agriculture and Food Topics, Chemistry Topics, Textile Topics, Environment and Sustainable Development Topics, Construction and Infrastructure Topics. The dominant field is *Chemistry*.

Distribution of technology fields in six-cluster set

Cluster 1; Information and Communication Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Agriculture and Food Topics. The dominant field is *Information and Communication*.

Cluster 2; Environment and Sustainable Development Topics, Energy and Natural Resources Topics, Construction and Infrastructure Topics, Machine and Material Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Chemistry Topics, Textile Topics, Agriculture and Food Topics. The dominant field is *Textile*.

Cluster 3; Agriculture and Food Topics, Textile Topics. The dominant field is *Agriculture and Food*

Cluster 4; Machine and Material Topics, Health and Medicine Topics, Chemistry Topics. The dominant field is *Health and Medicine*.

Cluster 5; Information and Communication Topics, Energy and Natural Resources Topics, Machine and Material Topics. The dominant field is *Machine and Materials*.

Cluster 6; Environment and Sustainable Development Topics, Agriculture and Food Topics. The dominant field is *Environment and Sustainable Development*.

Distribution of technology fields in five-cluster set

Cluster 1; Information and Communication Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Agriculture and Food Topics. The dominant field is *Information and Communication*.

Cluster 2; Environment and Sustainable Development Topics, Energy and Natural Resources Topics, Construction and Infrastructure Topics, Machine and Material Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Chemistry Topics, Textile Topics, Agriculture and Food Topics. The dominant field is *Textile*.

Cluster 3; Agriculture and Food Topics, Textile Topics. The dominant field is *Agriculture and Food*

Cluster 4; Machine and Material Topics, Health and Medicine Topics, Chemistry Topics. The dominant field is *Health and Medicine*.

Cluster 5; Information and Communication Topics, Energy and Natural Resources Topics, Machine and Material Topics. The dominant field is *Machine and Materials*.

Distribution of technology fields in four-cluster set

Cluster 1; Information and Communication Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Agriculture and Food Topics. The dominant field is *Information and Communication*.

Cluster 2; Environment and Sustainable Development Topics, Energy and Natural Resources Topics, Construction and Infrastructure Topics, Machine and Material Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Chemistry Topics, Textile Topics, Agriculture and Food Topics, Information and Communication Topics. The dominant field is *Machine and Materials.*

Cluster 3; Agriculture and Food Topics, Textile Topics. The dominant field is *Agriculture and Food*

Cluster 4; Machine and Material Topics, Health and Medicine Topics, Chemistry Topics. The dominant field is *Health and Medicine*.

Distribution of technology fields in three-cluster set

Cluster 1; Information and Communication Topics, Defense, Aviation and Space Industry Topics, Transportation and Tourism Topics, Agriculture and Food Topics. The dominant field is *Information and Communication*.

Cluster 2; Information and Communication Topics, Environment and Sustainable Development Topics, Energy and Natural Resources Topics, Construction and Infrastructure Topics, Machine and Material Topics, Health and Medicine Topics, Defense, Aviation and Space Industry Topics, Agriculture and Food Topics, Transportation and Tourism Topics, Chemistry Topics, Textile Topics. The dominant field is *Machine and Materials*.

Cluster 3; Agriculture and Food Topics, Textile Topics. The dominant field is *Agriculture and Food*