Christa Larsen, Marco Mevius, Jenny Kipper, Alfons Schmid (Eds.):
Information Systems for Regional Labour Market Monitoring. State of the Art and Perspectives

ISBN 978-3-86618-409-1, ISBN 978-3-86618-509-8 (e-book pdf),
Rainer Hampp Verlag, München u. Mering 2009, 224 S., € 27.80

Regional labour market monitoring based on web-based labour market information systems is becoming increasingly more common. This anthology shows the variety of such systems as implemented in various European countries. Two general types can be distinguished. The first are regionally-based information tools and secondly, national systems with regional sub-divisions. In addition, these show what information and communication technologies make possible, but they also give the first indications of where these technologies have their limits. This volume should trigger a wider discussion in this field.

This book was developed by several members of the European Network of Regional Labour Market Monitoring.

Key words: Regional Labour Market, Target Group Monitoring, Regional Information

Dr. Christa Larsen is a social scientist and Managing Director at the Institute for Economics, Labour and Culture in Frankfurt am Main. Her current work concentrates on regional labour market monitoring, regionalised analyses of labour markets for health workers, systems for information and diagnoses of labour markets and regional prognoses.

Dr. Marco Mevius is an expert for the management and implementation of large IT projects, designing new business processes and corresponding information in multidisciplinary environments. Currently he is Head of the division for Software Engineering (SE) and member of the Management Board at the Research Center for Information Technology (FZI) in Karlsruhe.

Jenny Kipper is a pedagogue and Scientist at the Institute for Economics, Labour and Culture (IWAK) in Frankfurt am Main. Her current activities are focussed on process of organisational learning, organisational development, and development of quality.

Dr. Alfons Schmid is Professor for Economics at the faculty of Social Sciences and Scientific Director of the Institute for Economics, Labour, and Culture (IWAK) at the Goethe-University in Frankfurt am Main (Germany). His main areas of research are regional developments of employment and labour market, new information technology and impacts on employment situation, regional competitiveness, attitudes towards welfare states, regional forecasting and regional studies in poverty and wealth.

Christa Larsen, Marco Mevius, Jenny Kipper, Alfons Schmid (Eds.)

Information Systems for Regional Labour Market Monitoring

State of the Art and Perspectives

Rainer Hampp Verlag München und Mering 2009

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

ISBN 978-3-86618-409-1 (print)
ISBN 978-3-86618-509-8 (e-book)
DOI 10.1688/9783866185098
1. Auflage, 2009

 Rainer Hampp Verlag München und Mering
Marktplatz 5 D – 86415 Mering

www.Hampp-Verlag.de

∞ *Dieses Buch ist auf säurefreiem und chlorfrei gebleichtem Papier gedruckt.*

Liebe Leserinnen und Leser!
Wir wollen Ihnen ein gutes Buch liefern. Wenn Sie aus irgendwelchen Gründen nicht zufrieden sind, wenden Sie sich bitte an uns.

Preface

The implementation and management of web-based information systems for the labour markets is a highly relevant topic from a practical point of view, because it offers many challenges to software engineers and economists. It represents an innovative application area for the concepts of information systems, which have been discussed in the computer science community for a long time.

Existing approaches to develop information systems for the labour markets are early attempts to provide solutions for the significant issues and problems, specific to this application area. The major difference between such systems and the more traditional applications lies in their management of information and content. While the developers of traditional information systems used to focus on defining adequate representations for the labour market data, the development of web-based information systems broadens this focus to content management, which includes: information modeling, authorship, volume, rate of change, and maintenance. The hyper-textual characteristic of web content makes it imperative that all stakeholders understand the nature of the information these applications deal with.

Today's web-based information systems for the labour markets are required to provide highly dynamic, personalised and interactive services, whose maintenance becomes anything but trivial, when the amount of relevant information exceeds a certain threshold. As the amount of labor market information grows, it may become difficult to adapt such a system to evolving requirements, which is why the two key non-functional requirements for such a system are from a software engineer's point of view: support for change and scalability.

As a result, the development of information systems for the labour markets is not a one-time activity, but a continuous process. Furthermore, this process should be controllable and traceable. Before setting out to implement such a system, the appropriate information sources must be identified in collaboration with domain experts and the type of required information must be specified. Traditionally, this has been done using questionnaires and interviews, both involving a lot of paper-work and a time-consuming manual evaluation of the collected data. Using these traditional approaches, it is not possible to promptly deliver critical information to specific target groups, which is why a new solution is needed, one which leverages the modern information and communication technologies of a web-based environment. Such a solution could be based on an integrated information portal, which incorporates several crucial functionalities to ensure valid and prompt information supply.

This book shows that an integrated labour market monitoring system, supporting branches and target group monitoring, can be significantly improved, if an adequate web-based information system is integrated in the monitoring processes. Moreover, the goal of this book is to sensitise actors in institutions and other interested parties to these aspects, and to encourage further research and development of such systems. This book provides an excellent introduction to the current state of the art and a comprehensive overview of the existing information systems for regional labour markets. The book covers many topics and links concepts to concrete applications, while at the same time it provides well-founded scientific approaches. All in all it is a superb compendium, recommended for all stakeholders interested in an effective monitoring of the labour markets, which will support institutions in addressing labor market monitoring in a more mature way.

Prof. Dr. Andreas Oberweis

Director of the Institute of Applied Informatics and Formal Description Methods (AIFB)

Contents

Introduction

1. Regional Labour Market Information Systems for Use in Monitoring

Exclamations such as “the information society has arrived on the labour market” are often not only referring to the fact that in recent years jobs in the Information and Communication Technology (ICT) area have appeared. Increasingly, more and more complex information on the labour market, employment and qualifications are brought together in web-based labour market information systems. These are not only used for internal business processes, but also made available to external customers. Often times, the information also is available at small regional levels, so that it can be used locally, at the regional or community level. Many of these regional labour market information systems are operated for the purpose of a continual regional labour market monitoring, or have been specifically built for this.

The utility of information and communication technologies for regional labour market monitoring is uncontroversial among labour market experts and researchers. This is above all based on the fact that regional labour market monitoring is conceptually strongly determined by the information needs of labour market actors (labour market and education politicians, associations, labour offices, educational and training institutions, employees, the unemployed etc.). The goal of the monitoring processes is to generate information adapted for the user and to communicate these appropriately. This permits the user, through the higher transparency, to make more informed decisions. Collectively, this result is an improvement of the functionality of the local labour markets. Critical to providing information to the user is that the user receives it quickly; using the least possible resources and that it is easily interpretable.

The traditional media for distributing information, such as printed studies or volumes of statistical tables, are unsatisfactory in meeting the conceptual requirements of monitoring. The data is frequently no longer current by the time it is printed. The work is not available everywhere or quickly when it is needed. The complex structure of the work prevents quickly finding the information required. In addition, if the user needs to re-use or re-work the data, new tables or structures have to be created. This takes resources and knowledge that the user, who often needs to decide quickly, does not have available. All these reasons indicate why such media are seldom, or not at all, used.

A similar situation arises with the challenges of using published official data. The selection of variables is frequently of limited use for the information needs of the user. Special analysis could be a response here. However, the user often lacks the time or material resources to create (or to have created) analyses to answer specific labour market relevant questions.

The limits of conventional media are, in practice, often surmounted by using a consultant or through surveys in which the focus has been placed on the required information. The difficulty here is that the material resources needed for a continuous observation of the regional labour market is not realistic for the majority of labour market actors. Thus, the labour market is only sporadically transparent. If, and to what degree, then, actions and decisions actually work is difficult to impossible to measure.

Through the use of information and communication technologies, these restrictions can be eliminated or significantly reduced. The data can be prepared according to a user profile and made available over the internet. Thus, the user has access to the information independent of location or time. This means that they can use it whenever they need it. In addition, many of the web-based labour market information systems offer simple and current data slices, which are clearly interpretable. The user can also download these tables or graphs from the internet and integrate these into their own documents. This way, the most current data is continuously available, without the user having to invest sizable resources of time and money. These are the real motivations to actually utilising data.

In the meantime, web-based labour market information systems have been implemented in several European countries, or are under construction, to enable a robust regional labour market monitoring. This book attempts to provide a look at the variety and diversity of such systems, and to show what efforts currently exist and where further developmental perspectives can be located.

2. Developmental Directions of Web-Based Labour Market Information Systems in various European Countries

The development of web-based labour market information systems is determined in individual European countries by various forces, which can be located within organisations, institutions or actors. An organisational context appears to be necessary when they have a long history in developing innovative public statistics, as will be shown in various examples in this book. In many cases, this seems to transition, more or less automatically, into a web-based information delivery. Very often, there is a purely statistical system built up at first, which however can feature here and there dynamic elements. Furthermore, it will become clear that institutional breaks can also work positively. This will be seen in the examples from some Eastern European countries, and in the case of a systematic change in controlling systems, such as the de-centralisation process in France. These “new starts” permit new creations to be directly built onto the current state of information and communication technologies. Thirdly, there are, above all, the decision making actors in organisations and politics who can trigger such developments. Their openness and affinity to technology work positively, as well as national culture and its corresponding experience pool in working with information and communication technologies.

The directions of development for the labour market information systems described in this book are classified along five dimensions.

Technical Dimension

Many of the labour market information systems are exclusively static systems. Within this group, the spectrum reaches from making available study reports to the possibility of downloading single graphs and tables in common formats. In between, there are mixed forms. These systems exist in most cases already for some time; often these static systems serve as the first step toward the development of a dynamic system. Dynamic systems are characterised here by having a database on the back end. In the simplest case, specific labour market data is first manually processed and then saved. These could be then activated through pre-defined templates and could be saved by the user. In most of these systems, a set of selectable characteristic combinations are provided. Notable on these systems is that the web interface must be very well structured in order to provide the user with a

quick and simple orientation to the available topics. In a few cases, more powerful data base systems are being developed and are just now being implemented. These data bases incorporate an automatic incorporation of data from various sources (a process called federation). These systems include processes for quality and plausibility controls, syntax standardisation and much more. After the data has been processed, it can be summoned exactly like in the simpler cases.

Based on the number of cases described in this book, it cannot be estimated whether, as many experts claim, that there is a shift from the static to the dynamic forms of labour market information systems. In this connection, it is also open as to whether or not the more complex dynamic systems will achieve a higher usage rate, and acceptance, from the relevant labour market actors. First experiences from the scientific labour market research indicate the utility of such systems. It is, to our knowledge, not yet examined, what the acceptance will be from other user groups.

Contents Dimension

Closely related to the technical requirements of the labour market information systems is the dimension of the data or information content. Many of the static systems are focused on specific themes, such as vocational and further training, on skilled labour or on labour market political target groups, such as migrants. For a more complex description of a regional labour market, static and dynamic systems are used. The static systems are applied primarily when there is a clear amount of data and often there is a long period between updates. Presumably in these cases, the effort for the manual processing of updates is less than the administration and further development of a complex database. However, information systems based on several data bases are also used in Europe. Partially, very complex and deeply integrated systems are used, especially in those cases in which a very broad spectrum of users has to be served by the monitoring information.

How complex the informational content in a labour market information system can be, while maintaining a good level of user acceptance, is a current question to be addressed in the evaluation of the systems. It can be assumed, based on first experiences and evaluation results, that single user groups have different limits with respect to complexity.

Dimension of Organisational Location

Two different types of organisationally locating the labour market information systems can be found in European countries. Firstly, many systems are directly found in public administration at the state and regional levels (ministries and their associated organisations, statistical offices and labour agencies). These systems are characterised by various data collections with high data quality and longitudinal data. The data collections can be represented over several area levels and deliver thusly regional information that can be directly compared with higher aggregated levels of data. A significant challenge with many of these systems remains that, although the different data collections are available in parallel, they are not connected. This means that the collections are not harmonised with each other and a direct connection between the variables is not possible. The second type of labour market information system is located in (regional) research institutions. One can often find regionalised data selections, wherein the data originated from public sources, but have been subjected to a specialised analysis. This standardises public data from various sources. This regional data can deliver therefore a more coherent situation analysis

than the regionalised data from the public administration. The labour market information systems from research institutions frequently contain primary data periodically collected. These systems in the second type create a direct content connection to the public data. A reversed data flow is occasionally present in some systems. An example of this is when official data is insufficient and a relevant study or survey result is referenced and linked.

Information and communication technologies fundamentally open possibilities for connections between official statistics and the data base systems of the regional research institutions. This would surely improve the actuality of available data and reduce the resources that have to be used for migrating data. This is, however, still restricted by current legal and organisational conditions.

User Dimension

The majority of labour market information systems are conceptually aligned with a supply orientation. Information to a thematic complex is offered through the internet. The user and his informational needs are not systematically considered in the thematic presentation of the data or their processing. In how far the offered information meets the actual need in these cases is hard to judge. Few labour market information systems, and these are currently exclusively the regionally located systems, are conceptually demand oriented. In these cases, the users within the labour market and their informational needs are known and the contents of the monitoring systems are adapted accordingly. Integral to these systems is a continuous evaluation of the available information's appropriateness. Mixed forms of these approaches can also be found. These are primarily supply oriented systems in conception, but allow a limited possibility of accessing a complex data base drawn from an individual information profile. These systems are still in a very early phase of implementation and it remains to be seen, which users actually take advantage of such offerings.

Within some labour market information systems, especially those based within regions, an interface between the available information and social networks are being worked on. Some of the articles here address this issue. How a functional interface can be constructed is still at this time wide open and requires further empirical experience.

Dimension Usefulness and Sustainability

Labour market information systems that are to be used for regional labour market monitoring should contribute to an informed decision process for labour market actors, improving the functioning of regional labour markets. To date, the systems are still in the development stage, which allows no assessment of this target. The few evaluations that have been undertaken, concentrate nearly always on the usage frequencies. Questions of usefulness are considered by many operators and initiators of these systems, but are only sporadically and unsystematically measured. Usefulness is meant here in the sense of the actual utility of the information for individual users or user groups to make informed decisions resulting from this information. How far information and communication technologies open there new and simple possibilities for evaluation remains to be seen. It appears that the proof of its usefulness could make a significant contribution to securing the sustainability of the labour market monitoring.

The dimensions sketched out here are handled in various manners in the web-based labour market information systems described below and can be applied across all case studies.

3. Goals and Contents of the Anthology

Using the labour market information systems implemented for regional labour market monitoring described in this book, the goal is to show how varied their structures can be in Europe and even within national borders. It should become clear how specific conditions can lead to specific systems. This anthology should also provide an overview of the developmental stages and directions in this field. For all those wanting to create or support a labour market information system, this should inspire. Beyond that, the contributions to this anthology should create a solid foundation to work within the framework of the European Network of Regional Labour Market Monitoring and, for those members of the network, to expand their expertise.

The structure of the anthology is oriented on the organisational location of the systems. In Chapter 1, the regional labour market information systems are introduced. The second chapter addresses national information systems with regional divisions which are operated by the labour administration and by statistical offices. In the final, third, chapter the various perspectives and considerations are given.

Chapter 1.1 focuses on regional labour market information systems that provide a comprehensive overview on regional labour markets. Using the example of the Federal Observatory of the Labour Market in Lombardy, Mario Mezzanzanica and Mattia Martini show what it looks like to go from a profound conceptualisation all the way to the implementation of a highly complex regional labour market information system. Wim Herremans and Marleen Jacobs describe how a less complex system, the Flemish Labour Accounts, was created. Here they also emphasis how data stores can be regionally connected. Yvette Grelet, Bernhard Hillau and Céline Vivent use two regions in France to describe how different regional and local systems can be, even when the general legal conditions are comparable. They focus as well on the interfaces between data and regional or community networks that come into sight as users. To conclude, Mareike Lemke and Kai Huter present the information system “Labour Market Statistics for the REGIO Lake Constance”. They demonstrate clearly what challenges one has to face when constructing a cross-border labour market statistic. The solutions that have found their way into this system are interesting and of significance for other cross-border labour market information systems.

In Chapter 1.2, the regional focus is retained. However, the labour market information systems presented here correspond to specific themes such as skilled labour and qualifications. The two first systems introduced were developed to proactively counter the skilled labour shortage in eastern Germany. Jan Ulatowski presents a static branch-related information system. In the following contribution Markus Höhne, Carsten Kampe, Daniel Porep and Anja Walter offer a labour market information system that is part of a regional construct, and thereby shows an example of how these central interfaces between data and actors can be realised. In the second part of the chapter, there are three articles which address the issue of qualifications. Christian Baden describes the framework and requirements for the creation of a labour market information system for the further education and training area in the federal state of Hessen [Germany]. The challenge for this static system lays less on the technical details as in the data stores, whose generation is burdened with severe

difficulties. Franz Clément and Adrien Thomas follow up with an impressive explanation of the border region between Luxembourg, Germany, France and Belgium, where such an information system is needed for securing generational replacements. They show how it was constructed and its historical antecedents. The final article in this thematic block is from Jiří Braňka and Michal Lapáček. They sketch out the beginnings of a complex and modern information system in the Czech Republic. Future skill needs are seen here as the primary informational need for an information system in the qualifications area. This article also provides the transition into the second chapter. This is due to its handling of a national system that can be regionalised in certain sections. Additional examples of this will be treated in Chapter 2.

In Chapter 2, it is shown through examples how national data stores can be broken down into smaller areas. At the beginning of this chapter are two articles from network members located in the directions of their respective national labour offices. Karin Horvath describes, using the Austrian labour information system "Qualimonitor", how this system was generated within a specific framework and how it could be further developed. The latter aspect is drawn with attention to the use of modern information and communication technologies. John McGrath and Caroline Shally outline how the national Forecasting System in Ireland was regionalised and what the drivers behind this were. The next thematic block is composed of three articles that focus on the regionalisation of data from Statistical Offices. Ivo Baštýř and Jana Vavrečková offer the possibilities in the Czech Republic and emphasis particularly using the system for a target group oriented monitoring. In comparison, Martine Mespoulet reports on the situation in France. Here there are entirely different conditions than in Eastern Europe. She focuses especially on the de-centralisation and its effect on the official statistics. As third in this group, Hélène Venningen serves up a glimpse from Austria of how regional labour market data from a Full Register Based Census can be obtained. The end of this Chapter is comprised of two contributions from scientists. Aleksander Surdej discusses, using Poland, what data stores are available, how far they reach and, above all, how to assess the data quality. A similar perspective has been chosen by Kezban Çelik and Sibel Kalaycıoğlu in sketching out the available data stores in the Turkish labour market. Here as well, the data quality is a central issue. This will make abundantly clear, that the generation of official data cannot be regarded as unproblematic across Europe. There is a need for action here that must be met before a systematic and reliable regional labour market monitoring can be pursued.

The final Chapter 3 turns to the subjects that are relevant for the further development of regional labour market information systems. Firstly, the information user is again brought into sight. Ronald McQuaid and Colin Lindsay report on a group of operative actors, unemployed workers, electronically accessing information from the local Labour Agencies in Scotland. With this concrete example, the opportunities and limits of the information and communication technologies are shown. The following article directs itself to the strategic labour market actors. Though the Spanish labour unions, Miquel Bernal demonstrates the central meaning ascribed to electronic labour market information. The third article has Marco Ricceri showing the national perspectives from the previous articles at the European level. He formulates the possibilities of regional labour market monitoring through a wider information and communication technology usage and suggests an organisational location. Additionally, he proposes how a labour market information system can help to overcome the current crisis. The conclusion of this anthology is an article from Alfons Schmid and Christa Larsen.

They attempt to derive, based on the contributions to this volume, in what areas the future activities of the European Networks of Regional Labour Market Monitoring could go to promote developments in labour market information systems in the direction of regional labour market monitoring.

4. European Network of Regional Labour Market Monitoring

This anthology represents the third volume from the European Network of Regional Labour Market Monitoring[1]. Once again all the authors are network members and their articles represent the current status of their activities. In this way, they provide a robust foundation for continued fruitful discourses within the network. As editors of this book, we thank all the authors for their commitment, and for the openness with which they approached their subjects. We would also like to thank them for the time and energy they committed outside the realm of their daily activities. We are quite excited that, thanks to the engaged network members, already for the third year in a row, we can publish an anthology from the network.

Christa Larsen

Marco Mevius

Jenny Kipper

Alfons Schmid

Frankfurt, August 2009

[1] Volume 1: Larsen / Mathejczyk / Schmid (2007) Monitoring of Regional Labour Markets in European States. Concepts – Experiences – Perspectives.

Volume 2: Larsen / Mathejczyk / Kipper / Schmid (2008) Target Group Monitoring in European Regions. Empirical Findings and Conceptual Approaches.

1. Regionally based Labour Market Information Systems

1.1 Comprehensive Labour Market Overview

The Federal Observatory of the Labour Market in Lombardy: Models and Methods for the Construction of a Statistical Information System for Data Analysis.

Mattia Martini / Mario Mezzanzanica

Abstract

This paper presents a summary of the methodological and technical path chosen for the study and realisation of a federal Observatory for the study of the labour market in Lombardy, with specific attention to two points: firstly, the creation of the Observatory's organisational model based on a network structure aiming at optimising competences, knowledge and the experience of the various public and private actors playing different roles in the Lombardy labour market; secondly, the optimisation of data collected and dealt with for different ends by the various public institutions and private companies that make up the federal network. The realisation, therefore, of a Statistical Informative System (SIS) to integrate such stock and flow data has significant importance. The work described comes from study and research by CRISP in recent years and the collaboration received from both local and central public administration[2].

1. The Observatory of the Labour Market in Lombardy

1.1 The Institutional Framework

The Lombardy Region has been part of the association “Four motors of Europe” since 1988 and is considered to be not only one of the most industrially developed areas but also one of the driving powers of the economy of the entire European Union. Following recent European indications (Commissione Europea, 2000) Lombardy has redefined labour market organisation and governance (I.R. 28/06 n.22), handing the responsibility of territorial active policies to the Regions and Provinces. The legislator of the reform conceived the regional Observatory of the labour market as a body for the gathering, updating and analysis of data and useful information to monitor, elaborate and analyse effectively the efficacy of employment policies, the educational system, professional training, further education and the regional labour market trends.

The planning of the Observatory brought forth its aims:

- to create a unique informative system for the sharing of information on the labour market held by public institutions and private companies;
- to study models of analysis of phenomena to optimise information putting it then at the disposition of the various stakeholders (policy makers, public institutions and private companies, individuals and research centres);

[2] CRISP is an Inter-University research centre on public services

- to plan a new organisational model - a federal Observatory - able to optimise the specific knowledge held by the institutions and companies working within the territory of the Lombardy Region.

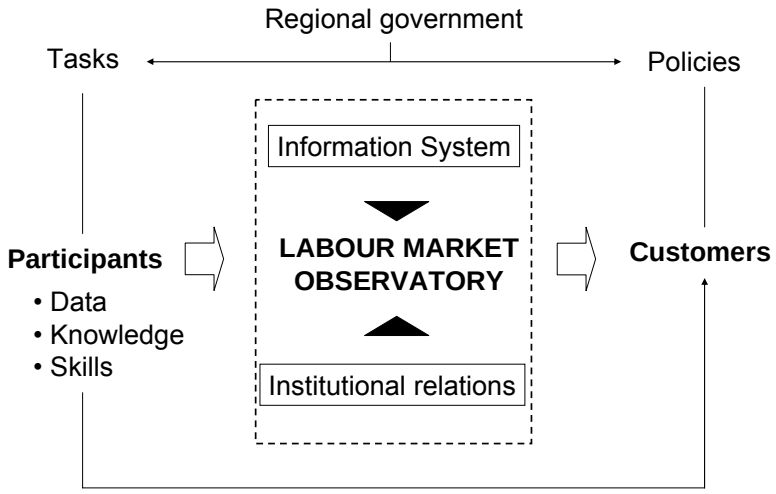
The regional Observatory of the labour market was founded in July 2008 following a regional legislative decree (D.g.r. 11 July 2008, n. 8/7605).

1.2 The Organisational Model and Services

The Observatory's organisational model was planned as a federal network, therefore, a collaborative network of public institutions and private companies competing within the labour market but contributing to the Observatory by sharing information and human resources and directly participating in the stages of planning and production of information providing services (Picherri, 2002).

Such interdependent collaboration aims at creating and therefore, having information of greater utility available to support one's own area of activity. The Observatory was planned as a system to create utility organised as a network (Allee, 2003; Normann and Ramirez, 1993). Figure 1 represents the Observatory's organisational model.

Figure 1: The regional Observatory federate network model for the Labour Market



The public institutions and private companies participating (provincial administrations, Chambers of Commerce, public and private employment services, INPS, INAIL)[3] provide relevant data and experience sharing, competences and human resources. A scientific board and technicians guarantee the internal governance of the Observatory's organisational model.
Experts, with particular technical and scientific experience in comparative processing and in-depth analysis of data provided by the member institutions and companies, make up the board. Technicians manage and update the Statistical Information System (paragraph 3) and all the activities concerning the realisation and diffusion of the results. The Region is responsible for external governance through the Agenzia Regionale Istruzione, Formazione e Lavoro (Regional Agency of Learning, Training and Employment), its correct functioning and the reaching of targets.

The Observatory provides two types of information services: one of an institutional nature and the other of an experimental nature. Firstly, the Observatory monitors the labour market both quarterly and annually presenting synthetic information derived

[3] INPS is the National Institute for Social Security; INAIL is the Italian Workers' Compensation Authority.

from the processing of data from the official statistics produced by the National Institute of Statistics (ISTAT) and from the processing of flow data from the appropriate private companies and public institutions. Secondly, the Observatory carries out specific group studies that can single out particular phenomena of special interest in this moment of change in the labour market[4].

The regional Observatory reveals its extremely important role in supporting regional and local policy makers in the stages of defining, actualising and evaluating active labour market policies. The data and the ensuing analyses by the Observatory are available to all those who need to know labour market trends and have also resulted useful in supporting the strategies and decisions of public and private operators in the employment and training system, in promoting and favouring the development of initiatives by the representative associations (the Chamber of Commerce, Unions, Associations representing work agencies), and in the orienteering and sustaining workers' career choices and employers' demand.

2. Analysis of the Labour Market by the Statistical Information System

A SIS can be defined as "an information System oriented to the gathering, memorization, transformation and distribution of statistical information (UNECE, 2000)[5]" which has the objective of regularly monitoring changes within the market in individual careers and in companies to intervene with effective active policies targeted at individuals in difficulty and requiring assistance, favoring their professional re-qualification and integration in the market itself.

The study of data from all the available sources provides complete and correct information and the definition of an integrated data model correlates the attributes from the distinct archives. Furthermore, recognition of the best procedures permits firstly, stability and secondly, automatisation of data provision and management of data processing.

Various projects have been sustained by a CRISP research group with participants from varying disciplines (systems of information processing, statistics, economy and management, law) concentrating on three major areas of study:

- an information system model;
- a definition of the methodologies of analysis;
- models to interpret and evaluate results.

[4] Examples concerning this second group of studies can be found in Mezzanzanica and Lovaglio (2008) and in Zavanella, Mezzanzanica, Pelagatti, Minotti and Martini (2008).

[5] The United Nations Economic Commission for Europe (UNECE).

2.1 The Model of the Statistical Information System

Organisations providing public services widely use information systems to support management of their operating processes. Such systems eventually produce enormous quantities of data that make up the patrimony of both single organisations and the community at large. The various archives store the descriptions of the different aspects and concepts of those areas receiving services from the institutions collecting and withholding specific data.

Therefore, the information scattered between the various databases has to be integrated into a single archive before deeper analysis of the phenomena associated with the concepts can be started which will then produce information and finally, knowledge from data. The high heterogeneity of the archives in question considering their structure, their subject matter, the data quality and the different levels of up-dating present the main problem addressed to integrate the data from the various archives and to analyse its statistical nature from a descriptive, explorative and previsional standpoint, as a support to decision making.

The methodology used to integrate the archives requires an accurate analysis of both the schema and subject matter of each database for the creation of a single integrated archive into which the various data will flow from the individual institutions in question. Models and techniques of analysis of the data quality are studied to identify the necessary upgrading so that the initial data reaches the necessary quality level for phenomenological analysis. Study continues with the design and consequently with the implementation of the architecture of the statistical information system based on data warehouse.
After this short introduction of the concepts of the integration of archives and data quality we can now briefly examine them.

2.1.1 Data Integration

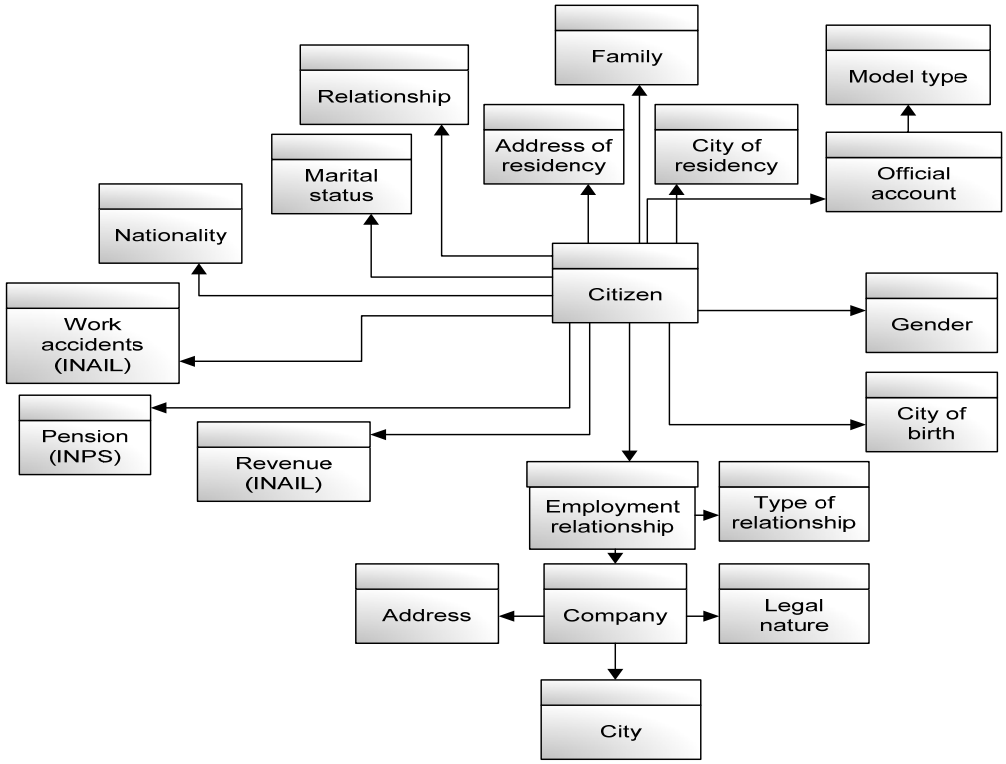
The initial stage of the integration of data schema consists in the analysis of the original administration archives to produce a description of the conceptual schema on which the archive was modeled and to normalise the individual original archives both structurally and subject matter wise[6].

The integrated data model planning aims at grouping the common information and systematically integrating the heterogeneous components pertaining to the various individual archives.

Figure 2 shows the conceptual schema of the personal data model obtained by the integration of information sources from public employment services (PES), the Revenue Office, the local Registry Office, INPS and INAIL.

[6] Structure: normalisation of the archive design requires a partial reorganisation of the logical structure by which data is organised to simplify the original archive design and above all to reduce information excess.Contents: normalisation of the archive contents consists in their standardisation and the conducting of possible variants of single data to a sole value.

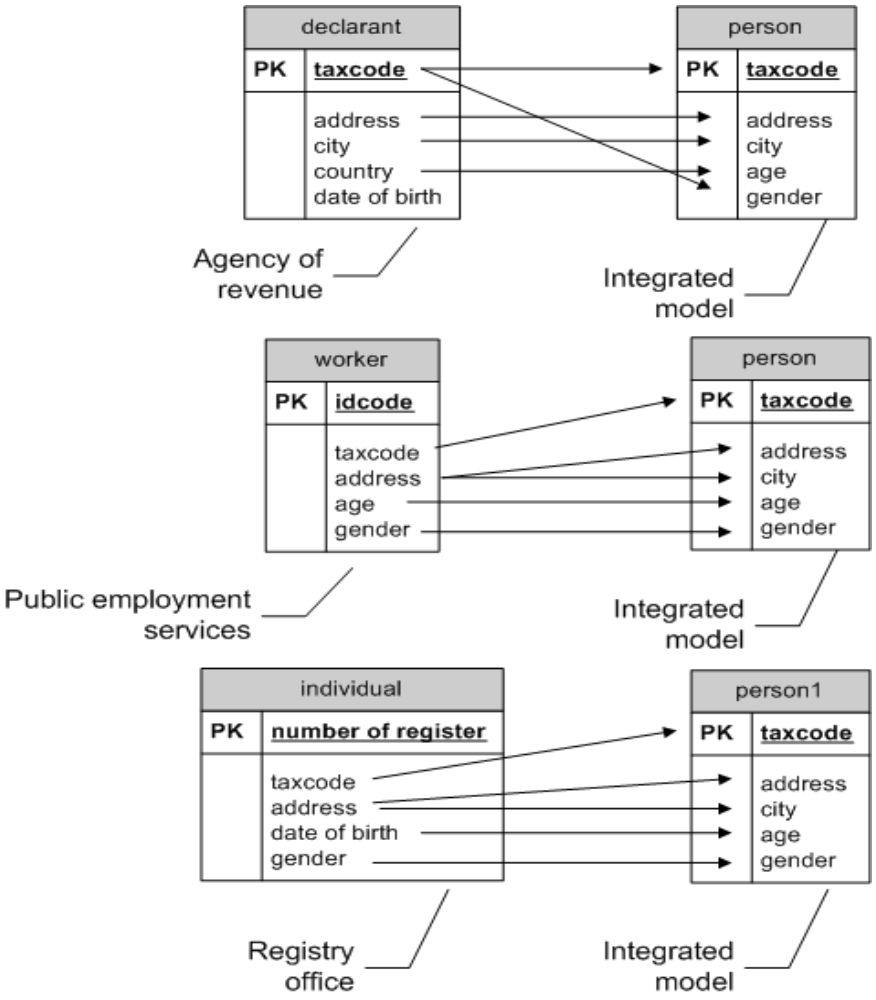
Figure 2: An integrated model



A later stage consists in the definition of the ‘mapping’ of both the database schema and instances, which will guide the data extraction from the normalised original archives and the following loading in the integrated archive.

Data schema mapping, which is an abstract activity, defines the similarities between the archives concentrating only on the design rather than details (e.g. the codes used). Figure 3 shows an example of the mapping of the entities ‘declarant’, ‘worker’ and ‘individual‘ versus the entity of the ‘person’ defined as an integrated data model (Mezzanzanica, 2009).

Figure 3: Individual entity maps vs a whole data model



2.1.2 Data Quality

The administrative archives are not usually of a sufficient quality to allow correct analysis; in fact errors and consistencies are found which can condition statistical analysis results. Therefore, quality of the archives must be checked by verifying:

- the conceptual model quality with which information is organised within the archive;
- the data value quality;
- the quality of data outflow.

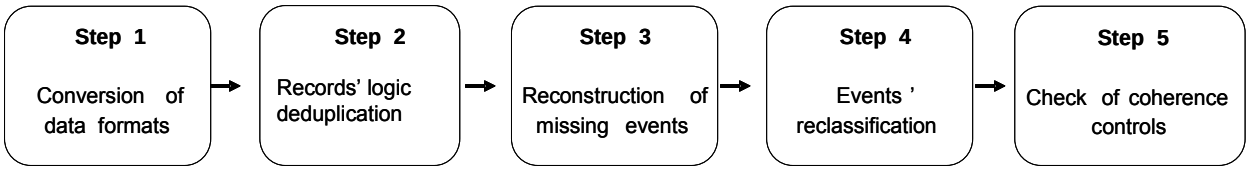
The conceptual quality and the data value quality are analysed according to the parameters stated in scientific literature: the *conceptual model for importance*, *aim*, *granularity attributes*; and the *data value for accuracy*, *updating*, *completeness* and *consistency* (Wang, 1998).

The final stage of verification and data upgrading concerns the reconstruction of data quality by techniques of database bashing and/or data edits (most used and consolidated methodologies).

We now present an example of the upgrading adopted for the construction of a SIS for the labour market in the Lombardy Region to exemplify the process. The final

table of the database analysed results in informative upgraded and homogeneous records of the subjects of interest.

Figure 4: The upgrading process



2.2 The Definition of the Methodologies of Analysis

A distinctive feature of a SIS for the labour market is the focus on the study of information processing models to single out innovative methodologies to represent regional labour market characteristics. Particular attention is given to longitudinal analysis considering the recent debates concerning accountability for those agencies providing services to the public. Such evaluation following a longitudinal logic permits an accurate reading of phenomena concerning efficiency and efficacy. The availability of information sources which are flows that summarise and provide ready information on the status and the whole chain of changes of the population observed, rather than surveys which are forecasted, brings novelty and propositional strength. The data within a SIS can be analysed using three different types of analysis:

- descriptive analysis (horizontal: the whole population in the period examined);
- advanced analysis (horizontal: the study of the population by the application of multivariate clustering modes);
- analysis ad hoc (vertical: referring to portions, subsections of the population examined) and synthetic indicators.

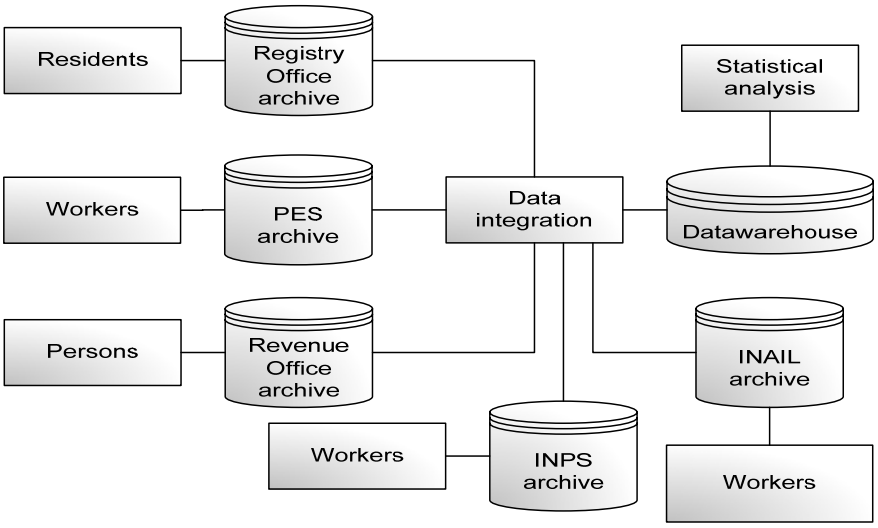
Analysis by clustering techniques, for example, has allowed the classification of work paths based on the types and duration of employment contracts; the evaluation of the temporal evolution of work experience; the grouping of homogenous subjects considering steps in contract types, thereby identifying the description of workers who keep one type of contract and the description of the bettering or worsening in terms of persistence in a status. The representation of the work paths of the entire population examined allows the analysis of the real trends during the period of analysis and therefore, increased knowledge for decisional support.

Finally, ad hoc analysis on a portion of the population concerned allows the study of specific targets: temporary workers, graduates, women.

2.3 Information Sources and Intervention Level

The data contained in a Statistical Information System can be used to efficiently monitor the evolution of local territorial phenomena and to answer effectively the concerns that emerge. Figure 5 shows the main information sources used for the construction of the SIS for the Regional labour market.

Figure 5: The Information Sources of the SIS for the labour market



Data from the local Registry Office and the Revenue Office is updated annually; INPS and INAIL provide information on an annual basis following a request and depending on availability, whereas PES provide mandatory data quarterly.

PES represent the primary data source concerning mandatory data (gathered telematically since January 1 2008). Such administration sources allow the production of structural statistics (describing the structure of a population or the characteristic elements of a company) and statistics of change (longitudinal analyses). The latter can be produced because administration data is periodically updated and provides information on the characteristics of the worker (contract duration, type of contract, qualifications, economic sector, etc.) such information, rather than the traditional stock information, allows a correct definition of employment trends. Furthermore, given that data is gathered continually, a unit can be followed in time therefore allowing flow analysis over a long time span (Mezzanzanica and Lovaglio, 2008).

2.4 The Architecture of the Statistical Information System

The finding of technological solutions able to offer a lasting high level of configurabilty, reliability and flexibility allows the Statistical Information System to respond fully to the needs expressed. This paragraph describes the architectural solution studied and realised. The solution positioned in the ambit of 'Business Intelligence' is made up from a Data Warehouse of software modules realised in an open source ambit.

2.4.1 Architecture of the data warehouse

The combining of the heterogeneous characteristics of data contributing to the warehouse while guaranteeing best performance and information availability make up one of the distinctive features of data warehouse planning.

Figure 6: Data warehouse architecture

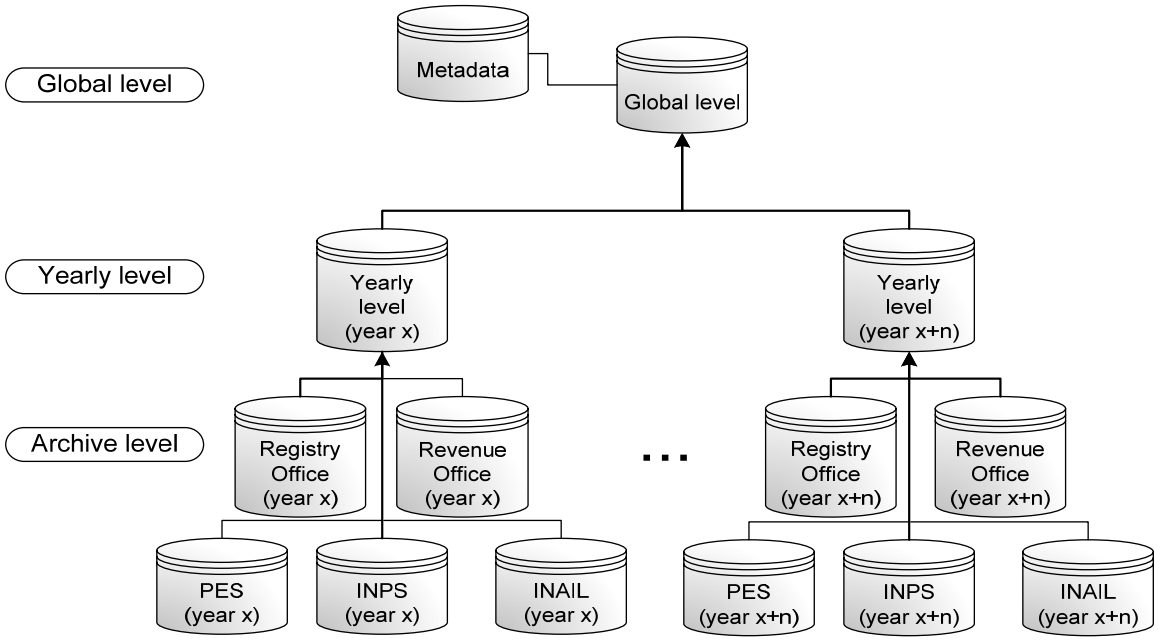


Figure 6 shows how the data warehouse is planned on 3 distinct levels to reach the fore said objectives:

- a level containing the single deliveries (administration sources) loaded in the data warehouse that can be managed, treated and analysed individually;
- a yearly level in which the first level of data integration is reached subdividing the information according to the period of observation; this level holds both flow and stock information concerning the year of interest;
- a global level that integrates all the available information and offers a complete vision of the whole database.

To respond to specific needs of analysis, some 'data mart' have also been realised which have subsections of information of particular interest that allow the selection of information according to the period of observation or according to particular characteristics requested or which are still undergoing particular treatment (e.g. simulations).