Business Intelligence in Healthcare - Data Mining Techniques as a Possible Hospital Management Tool in Austria

Master Thesis

In partial fulfilment of the requirements
for the degree

,,Master of Arts (MA)"

Master Program
,,International Health and Social Management"
Management Center Innsbruck

Supervisor:
FH-Prof. Dr. Nils Mevenkamp

Author:
Marten Schmied
1410360006

Declaration in Lieu of Oath

I hereby declare, under oath, that this master thesis has been my independent work and has not been aided with any prohibited means. I declare, to the best of my knowledge and belief, that all passages taken from published and unpublished sources or documents have been reproduced whether as original, slightly changed or in thought, have been mentioned as such at the corresponding places of the thesis, by citation, where the extent of the original quotes is indicated.

The paper has not been submitted for evaluation to another examination authority or has been published in this form or another

Prague, 30th of June 2016

Marten Schmied

Acknowledgment

"Gratitude is not only the greatest of virtues, but the parent of all others."

Cicero

A number of people helped me through the process of writing this thesis by their support, professional input, motivation or mere presence and I shall not forget to express my gratitude for all the backing I received. Particularly my dear Lisa, who had to take the bulk of my mood swings related to the ups and downs of this work. I would like to thank my family for their understanding even though my worries often were foreign to them. Without my supervisor Prof. Mevenkamp's advice I may have been lost a few times and I am grateful for his patience and the reassurance he provided. Thank you to all the Interview Partners taking their time to assist with what from their perspective must have been very basic knowledge. Also my dear friend Nico helped a great deal with sharing his experiences in computer science. Thank you to my friends Liz and George, Toko, Griese, Hannes and Caro for remaining my friends even though I have been fairly caught up in this work for the last half a year. Finally it is Adam whom I have to thank for taking my mind of writing a thesis once in a while.

Table of Content

List of Abbreviations

AI	- Artificial Intelligence
ANN	- Artificial Neural Network
BI	- Business Intelligence
CIS	- Clinical Information System(s)
CRM	- Customer Relations Management
DB	- Database
DBMS	- Database Management System
DSS	- Decision Support Systems
DW	- Data Warehouse
ECM	- Enterprise Content Management
EHR	- Electronic Health Record
EIS	- Enterprise Information System
ERP	- Enterprise Resource Planning
ESS	- Executive Support Systems
ETL	- Extract, Load, Transform
HIS	- Hospital Information System(s)
HIT	- Health Information Technology
HL7	- Health Level Seven
HR	- Human Resources
KDD	- Knowledge Discovery in Databases
MIS	- Management Information Systems
OLAP	- On-line analytical processing
RFID	- Radio Frequency Identification
SCM	- Supply Chain Management
SVM	- Support Vector Machines

List of Figures

List of Tables

1. Introduction

It is a widely acknowledged fact that public healthcare systems feel an increasing need for economic planning. Gaining insight into the processes and factors influencing a company's efficiency can be achieved by a variety of managerial tools. Analysis of business related data and the according tools commonly get referred to as Business Intelligence (BI). The art of BI, traditionally performed by a small clique of "data experts" that had the knowledge and skills to automatically search data and discover unseen patterns (Jing-song et al. 2011, p. 159) is now a service offered by an array of consulting companies and became a natural part of most industries. Despite the fact that all practiced tools aim to enhance a company's economic performance by structuring and further analysis of admissible data, the capability of a BI solution critically differs with its depth and functionality.

According to Gluchowski et al. (2008, p. 90) Online-Analytical-Processing (OLAP), Management Information Systems (MIS), Decision Support Systems (DSS) and Executive Support Systems (ESS) compress and present business relevant data serving the classical functions of BI. An expanded operation includes a more analytical approach using tools such as Key Performance Indicators, Balanced Score Cards and Data Mining. Finally it is useful to include the Data Warehouse (DW) into the widest definition of BI (Gluchowski et al. 2001), p.7) as without the adequate formatting and storage of data none of the above functions could be exercised.

So the complexity of BI can greatly vary from database management systems (DBMS) eventually including automated reports and functions of decision support to highly sophisticated analytical software extracting new knowledge from unseen patterns. This last process using refined algorithms based on artificial neural networks (ANN) or machine-learning techniques is commonly referred to as Knowledge Discovery in Databases (KDD) and may be applied to a variety of situations (DeGruy 2000, 62ff). *Data-Mining* (DM) is the analytical step in a KDD process and preferably done on big homogenous data sets.

By now these techniques have also descended into the healthcare sector, with KDD applications being found in the evaluation of treatment effectiveness, the management of healthcare, customer relationship management and the detection of fraud and abuse (Koh, Tan 2005, 66ff). Further possibilities for the use of DM are describe within System Biology, Hospital

Management and the Medical Device Industry including the Pharmaceutical Industry (Durairaj, Ranjani 2013, 31f).

A particular difficulty with DM applications is the huge amount of healthcare related data that has to be modified and organized in order to be mineable. Approximately four terabytes of raw sequencing data may be expected from each individual and healthcare provider hold data matrixes with up to hundreds of thousands of patients (Chen et al. 2012, p. 1171). So without adequate tools for analysis all stakeholders in the healthcare "can still be rich in data but poor in information" (Olszak, Batko 2012, p. 976).

Unfortunately it is particularly challenging to analyse business processes in healthcare organisations as the underlying processes are of "highly dynamic, complex, ad hoc, and multi-disciplinary nature" (Rebuge, Ferreira 2012, p. 99). Yet potential benefits from the use of KDD software as a tool in hospital management get described in the literature. Possible benefits are manifold: the frequency with which operating theatres could be used may be improved, the allocation of material and staff could be optimized, patient clusters and the impact of certain diseases on the hospitals services identified (Alapont et al. 2005). DM may also establish models capable of determining whether or not the purchase of expensive diagnostic machinery, such as a MRT, is economical (Raphael 2014, 41f; Kurbjuhn, B., Schult, R. 2010).

With a first standard for DM products already established in 1999 (DeGruy 2000, p. 62) and a general similarity in the structure of most Hospital Information Systems Alapont et al. (2005) argue that it is possible to construct a tool that automates many processes for DM that could easily be exported to a string of hospitals. An expanding accessibility of specialized DM software together with the healthcare sector being a rather new field for the application of such (Olszak, Batko 2012, p. 969) it seems very probable that the full potential of these techniques for the hospital management has yet to be utilized.

2. Goal & Research Question

Due to economic restraints and the increasing pressure to make hospitals more cost efficient the successful application of DM is offering a set of possibilities for improving the operational functioning and the strategic planning. Moreover these techniques are about to be increasingly user-friendly and therefore applied with smaller effort and expenses.

BI solutions for hospitals get offered by a series of companies, ranging from big international consulting groups such as Gartner and Siemens Healthcare to smaller enterprises offering individually programmed systems like the Healthcare Explorer from the TIP Group in Graz. At least 160 providers of BI work within hospitals (Breitschwerdt 2011, p. 44), thus making it difficult to assess the quality of all obtainable solutions. Furthermore it is unknown if sophisticated strategies for the usage of Information Technology (IT) generating BI in Austrian hospitals do exist and to which extend they get implemented. Out of 27 healthcare consulting corporations in Austria only six can be identified to focus on Information & Communication technologies (Qualitas 2009). As the process of KDD requires a certain dedication to data management and thorough knowledge of (health) informatics the list of possible actors assisting with DM projects for hospitals could very well be extend by individual programmers not related to the hospital management particularly.

Individual, do-it-yourself BI solutions could be a very cost-saving alternative to commercially systems thanks to freely available, so-called *Open Source* software for DM that generally could be applied by any IT technician possessing the skills, administrative rights and access to mineable data. Popular free modules for DM are the WEKA package in Java script or Orange for Python. A number of pre-set DM applications also exist within the programming language R or the ProM platform. Thus far it has been a problem that these free programs' functionality seems somewhat limited compared to the commercial tools (Thomsen, Pedersen 2008, p. 1). Chen et al. (2012, p. 1171) describe that Miller (2012) found shortcomings within the analytics of health related big data compared to BI applications in e-commerce as they "rarely taken advantage of scalable analytical methods or computational platforms".

While it becomes clear that healthcare seems to be rather underdeveloped regarding BI access to studies on the tools used for hospital management or the quality of data processed, stored and ready for business related analysis seems very limited. Furthermore no study appears to establish which BI functions are already facilitated in Austrian hospitals and therefore it is unfeasible to estimate the possibilities that do not get utilized. That is why this thesis aims to answer the following questions:

> *"What kind of data is electronically processed in Austrian hospitals and to what extend does this data get processed in terms of BI?*
>
> *And:*
>
> *Is it possible to identify unused capacities for DM in Austrian hospitals?"*

The thesis describes the current use of BI in Austrian hospitals and investigates possible procedures to further introduce DM software into Hospital Management. It is to establish the level on which data gets generated, stored and processed within this environment in Austria. Finally the goal was to make recommendations on how software-based analytical techniques can be best applied within the countries' public hospitals. The thesis' topic combines high potential benefits from successfully implementing DM into Hospital Management with the increasing need for efficient and sustainable planning in health care.

3. Methodology

The first part of this thesis gives the theoretical background on data processing in general public hospitals. Chapter 4 explains the data that is generated by the different organisational systems within a general hospital. According to a processing model the data sources are examined, it is shown how data gets handled and stored and how it may then be used to gain insights in terms of BI. Chapter 5 ensures a basic understanding of the KDD process and DM as its essential tool. Different DM techniques are introduced along their most frequent applications in healthcare. The theory concludes with determining the categories on which relevant literature concerning the research goals were identified.

A mixed method was used for the thesis' empirical part. Findings stem from the literature review and four semi structured expert interviews and were divided into the assessment of HIS on one side and the identification of possible DM tools for the Hospital management on the other.

A detailed listing of the search results for the assessment of HIS is attached in Annex A. Notably some of the searches yielded high numbers of results but only included very few relevant papers. Extending the search from Austria to Germany in PubMed, for example, gave seventeen results half of which were dealing with one particular German clinic between 1994 and 1995. Finding difficulties in the definition of terms and functionalities, Chapter 5.3 gives further details on how relevant literature on the use of DM Applications specifically for Hospital Management was identified. Search results are also summarized in Annex A.

Semi structured Expert Interviews were conducted with the interviewees summarized in Table 1, they were all identified as experts on at least one of the thesis's aspects. The interview partners were initially contacted via e-mail; interviews were conducted via Skype and lasted between 30 and 45 minutes. All participants were called on landlines during their working hours. Consent was granted at the beginning of each Interview. The interviews were recorded and those passages relevant to the thesis' research questions transcribed for quotation. An exemplary interview guideline is included in Annex B, the partial transcripts in Annex C.

Inter-viewee	Name	Position	Organisation
I#1	Fr. Univ.-Prof. Dr. Elske Ammenwerth	•Professor for Health Informatics •Head of the Institute for Biomedical Informatics •Member of multiple professional associations and editorial boards	UMIT - University for Health Sciences, Medical Informatics and Technology
I#2	Herr Christian Stark	•Project Manager Clinical Information Systeme •Coordinator Implementation of ELGA domain for Tyrol	Tirol Kliniken GmbH (TILAK)
I#3	Herr Karl Kocever	•Head of the Information and Communication Department •Managing Director Steinerärkische Medizinarchiv GesmbH (marc)	Steiermärkische Krankenanstaltengesellschaft m.b.H. (KAGes)
I#4	Two interviewees from a fourth interview wished to remain anonymous.	•Head of the Software Engeneering Department	Anonymous
I#5		•Head of the statistical department	

Table 1: Partners for Expert Interview

Findings from the interviews as much as quotations from German literature were translated by the author of the thesis himself.

Chapter 6 finally analyses the relevant literature and interviews in order to assess IT penetration and technological structure in Austrian public hospitals. Simultaneously an adequate tool to measure the functionality of HIS in terms of extracting information relevant to the management as in generating BI was searched for. The chapter's second part summarizes the studies found to be relevant regarding the application of DM techniques on HIS in order to identify feasible models and their value for individual hospitals or networks. Limitations to this thesis and an outlook over future applications and research are discussed in Chapter 7.

The other research questions were answered by estimating the IT capabilities of Austrian public hospitals on the one side and grant overview over possible DM applications for the hospital management on the other, providing a contemporary report on this topic. The thesis contributes to understanding the status quo of Austrian hospital IT and the unutilized potential of DM as a hospital management tool. Besides providing exploratory work for future studies to further define and optimize BI facilitation in this sector, the report helps to make recommendations for future practices.

4. Data Processing in Hospitals and the Use of Business Intelligence

The provision of public healthcare services is a very particular, in economic regards an *imperfect market* with special conditions (Mwachofi, Al-Assaf 2011), that generates huge amounts of complex data. A general model for handling data towards analysis, visualized in Figure 1, was used to explain the aspects related to this process.

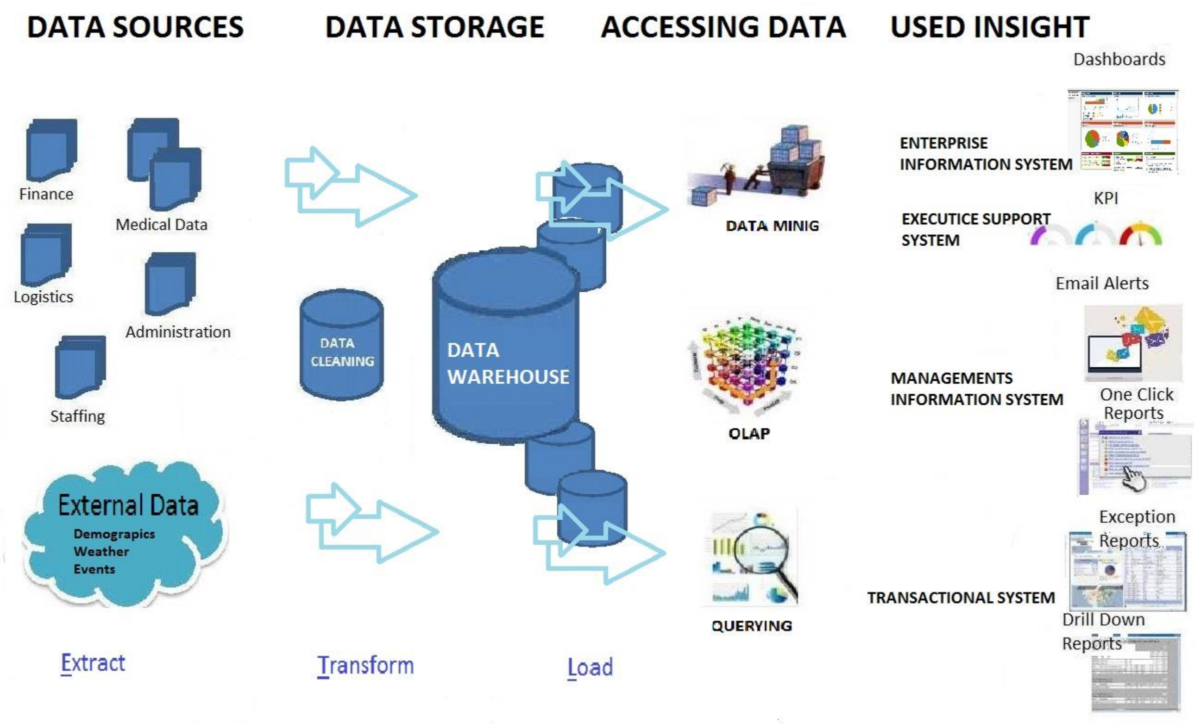

Figure 1: Data Handling Towards Analysis
(adapted from: Hao-Yun et al. 2016; Olszak, Batko 2012, p. 973; Raphael 2014, p. 20)

This chapter enables the reader to understand the components, structural basics and functionalities of Hospital Information Systems (HIS).

4.1 Data Sources

Within this thesis the term *Data* is to be interpreted very broadly. Data is anything recorded or captured disregarding the format; data may be made up from text, calculations such as in Excel files or any kind of image. Obviously the question whether or not data is stored electronically is

of great importance for this topic. In clinical practise the attempt to decipher a physician's handwriting may lead to misinterpretation of data and the *paperless hospital* as much as *eHealth* are topics related to this thesis as its focus is on IT-based analyses. Information is knowledge extractable from all data if rightly understood. For this setting it means that when a doctor tells a nurse about a patients' diagnosis or a nurse calls the diagnostics department in order to arrange a consultation, it is information that does get transmitted without being reflected in the data until the according patient records or schedules are changed accordingly.
All analysis of business related information in order to guide and support managerial decision making is based on interpreting raw data. Within the healthcare sector an abundance of available data is generated and stored. Jing-Song et. al (2011, p. 147) speak of gigabytes per hour volume. The modules handling, storing and distributing patient related data within a hospital is commonly referred to as Hospital Information System (HIS) or understood as the use of Health Information Technology (HIT). Tsumoto & Hirao (n.d., p. 2) describe the HIS as "cyberspace of medical orders" due to their core function of transmitting orders between the different hospital departments.

As HIS hold sensitive data, such as personal information and diagnoses, the protection of this data and privacy concerns are of utmost importance. Sharing a patients confidential information in a widespread healthcare environment may compromise the necessary privacy (Parvez et al. 2015, p. 46).
In Austria all electronic health data is subject to legislation according to the *Gesundheitstelematikgesetz* (GTelG) from 1st of January 2005 and its December 2012 revision. This for the thesis principle law regulates standards and procedures for the storage and transfer of health related data, like mandatory encryption. Further laws concerning patient related information or the handling of data within a medical environment are the legal regulations for medical practitioners (Ärztegesetz), nurses (Gesundheits- und Krankenpflege-gesetz), medical technical services (Medizinisch-technische Dienste Gesetz), the legislation on public hospital and rehabilitation services (Krankenanstalten und Kuranstalten Gesetz) that may be extended and specialised by state specific regulations and the general law on data security (Datenschutgesetz).

A HIS serves administrative and clinical tasks and reflecting these multi-disciplinary characteristics is compiled from a number of components:

1. Clinical information system (CIS) or dynamic data

From the medical point of view a hospital's CIS contains the most essential information about a patient. The CIS can be understood as repository of all clinical data regarding the patients' treatment. Dynamic or medical data is not only composed of the doctors' and nurses' documentation but of all other reports, including diagnostics like lab results and image files. Due to the conceivably high number of involved actors and changes in a patient's condition during hospitalization, the medical data may be adapted and extended on a frequent base and by a broad variety of contributors. The sheer amount of clinical data implies a high possible value for analyses. Particularly, for the improvement of therapies but attributable to the heterogeneity of the files (tables, text and images) this information can be hard to extrapolate. The Health Level Seven (HL7) is a format that "may be seen as a request form to access specific information related to a patient" and severs as "a widely used standard for any health related data" (O'Neill 2011, p. 37). Moreover for the handling and sharing of digital medical images the Digital Imaging and Communications in Medicine has become the DICOM standard and is extensively used within the field of radiology (O'Neill 2011, p. 39). It is specifically the hospitals' imaging departments that often show well established and firm IT subsystems in the form of a Radiology Information Systems (RIS) and Picture Archiving and Communications Systems (PACS).

2. Core or administrative Data

This data may also be described as a patient's administrative data and contains the patient's home address, the insurance provided and diagnosis coded in ICD-10. (Tsumoto, Hirano n.d., p. 2). This information may be transferred from the patient to the service provider using an electronic health record (EHR). In Austria a nationwide and dedicated use of EHR is underway with the *ELGA* project, already being partially implemented and about to be fully introduced during 2017. Administrative or core data holds information that is regularly exchanged between departments and wards, e.g. readout of e-card or a patient bringing their file to the ward. Changes to the data are rather infrequent and initiated by only very few contributors like the patient, doctors or the insurance company. For hospital services this information gets mostly read in order to verify a patient's identity and spread the according messages between different units when the patient gets referred. Within an HIS it is customary to establish individual patients' ID to assign case numbers, avoid confusion and locate the person. Interfaces to other IT subsystems exist to the controlling as the billing is based on the patient's admission dates and diagnoses. The increasing amount of standardized software for the administrative data like SAP R/3 means more interfaces between HIS and CIS, particularly regarding medication and consumption statistics (Raphael 2014, p. 22).

Data that is not related to the patient directly but handled within a HIS can be organized into three categories:

3. Controlling, Accounting

Controlling covers the financial sector of a hospital and employees its own specialists. The Austrian DRG System, the *Leistungsorientierte Krankenanstaltenfinanzierung*, is used for the reimbursement of hospitals according to treated cases (Potocnik 2006, p. 3). The most important economic indicators regarding a hospital's operational functioning could be summarized by reports based on the financial information processed in the controlling and accounting department. Furthermore it is the financing department's data that strongly correlates with information from all other subsystems. Patient's core and dynamic data is needed for billing the insurance companies according to the services provided to any given patient, while for accounting purposes also basic information from logistics and the Human Resources (HR) department has to be provided.

4. Logistics

Logistics' data is mostly concerned with running the hospital's inventory. Supply Chain Management (SCM) became an increasingly important concept within hospitals, particularly as dramatic inefficiencies and years of adversarial relations to suppliers have to be overcome (Toba et al., 2006, p. 49). To positively impact bottom line profitability the hospital's internal ordering systems are usually supported by IT through their own Enterprise Resource Planning (ERP) system like SAP MM or AMOR (Raphael 2014, p. 199). The pharmacy, providing an essential part of the hospital's medical services, has an important role within logistics and is usually managed through an individual system. Due to the frequency with which medication is changed while a patient remains hospitalized, a closely knitted interface between patient's dynamic data and the pharmacy's system may prove beneficial and reduce waste (Tolan et al. 2015). Logistic's data is of a particular interest regarding business analysis when inventory can be linked to specific procedures or patients. Implants and medicine coated stents, for example, could be traced down to a specific patient receiving the item, while a lot of materials could only be attributed to a certain ward ordering them. Moreover the use of pre-fabricated packages, e.g. a catheterization kit, makes it possible to allocate some material to a certain procedure.

5. Staffing

Like the patient related data also information about staff can be divided into a core and a dynamic component. The core contains the employee's basic biography, e.g. a copy of the CV or the length of employment with the company. Also holding personal information like age and maternal status, this data is subject to certain precautions regarding data safety and gets transferred to another subsystem just in an anonymous, aggregated form (Raphael 2014, p. 198). Staff's dynamic data in terms of which tasks have been accomplished may be a legal requirement to document who is responsible for a certain activity involving the patient. This information is usually resembled in the medical documentation and operational reports.

Finally it is possible to identify data holding valuable information but being generated outside the hospital. Such data may be readily available in the form of governmental reports and statistics and could be added to the internal analysis.

6. External Data

External information on local demographics as much as studies on disease prevalence within the region may help to foster precise predictions regarding the flow of patients. Weather may impact average admissions as simply speaking with frost more people break their legs and in hot weather an increased number of patients get hospitalized with dehydration and heat stroke. For the same reason the social calendar may have some significant correlation to the number of admissions as more accidents happen around public holidays and important events (Alapont et al. 2005, p. 4). Publically available information about other healthcare providers may also help to analyse the market situation and enhance strategical planning (Raphael 2014, 201f).

As external data does not naturally get processed within a hospital, to include them into analytics does require some dedication to data preparation and model building.

4.2 Data Storage

A vital process in data management, particularly with data warehousing is abbreviated with ETL standing for: Extract, Transform and Load. These three steps have to be conducted to introduce raw data into an existing system or merge files from different subsystems into a common warehousing solution. ETL serves as an industry standard for data processing and effectively is a pathway for linking different datasets (Sakshi 2012).

Each dataset may consolidate different formats such as flat files or xml. While these formats might also be organized in diverse data structures the extraction phase generally aims to homogenize all data into one format fit for transformation. To achieve this goal data validation is a vital part conducted on extraction. While the data is pulled from its origin, the included values do get checked against a set of rules. Upon rejection the system may report back the invalid data to the administrators.

Transformation sets the specifics for the chosen format and cleans the extracted data of invalid or missing parameters often referred to as *Noise*. While some of the data may not need transformation and can directly pass through to the server, common tasks performed during transformation are:

- Selection of the attributes or certain columns needed for further processing
- Translation into consistent values, e.g. M for Male throughout all data
- Order data to improve performance of searches
- Merging information from multiple sources
- Delete double values or repetitions
- Data aggregation
- Assign unique identifiers to different objects in the database (Surrogat-key)
- Turning multiple columns into multiple rows (Transposing) or vice versa (Pivoting)
- Splitting single into multiple columns

(Menken 2013, p. 45; Sakshi 2012)

When the data is readily extracted and meets the serves specifications it gets loaded onto the main server. From here the data can be moved to smaller subsets, the dart marts which are accessible for further processing. Data bases as a first step and then data warehouses build from them are rapidly emerging as they become easier in handling. Nowadays not solemnly IT departments work on them but business analysts themselves may structure them according to their specific needs (DeGruy 2000, p. 60).The Data Warehouse (DW) as the main storage of all data has its own architecture so that the information can be further used and analysed without doing alterations to the original source. By providing logically centralized data the warehouse enables users to quickly access information leading to increasing efficiency in inquiries. Despite a vast amount of literature about the structural design of data warehouses no widely accepted standard for its modelling has been accepted yet (Avin 2008, p. 16). The intervals in which information in the DW gets updated and overwritten can be programmed to fit the companies' needs and may vary from an hourly loading of new data to historical storage for a one year period. ETL is commonly executed on all its three levels simultaneously. While the extraction continues readily pulled data gets transformed and loading of the data begins before the other steps are finalized. Lloyd (2011, p. 53) argues that ETL tools are to be seen as its own "piece of the business intelligence puzzle" not a mere necessity for maintaining a DW. For this thesis a DW and the accompanied processes in data management were acknowledged as a prerequisite to efficiently use data for BI. The value of using this data is constituted by the generation of accessible and adequate reports and predictions. Raphael (2014, p. 26) points out that in most German hospitals no data warehouses exist yet. If this situation is assumed to be similar in Austria, attempts to generate BI would seriously be hindered by lacking the necessary infrastructure in IT.

The holistic approach to integrated and centralized data storage within a company is commonly referred to as Enterprise Content Management (ECM). An example for the complexity of such structures is shown in Figure 2 (AGFA), presenting the Belgium Company's Agfa Healthcare's ECM solution for hospital's called HYDMedia G5. A high number of interfaces resulting from the six defined workflow levels with Agfa's G5 Services as the central connector catering to the different modalities of data storage and the clients are visible. While this showcases a commercial solution offered by a private company, any hospital or federation of clinics may have an individual IT structure with limited interfaces between the different departments.

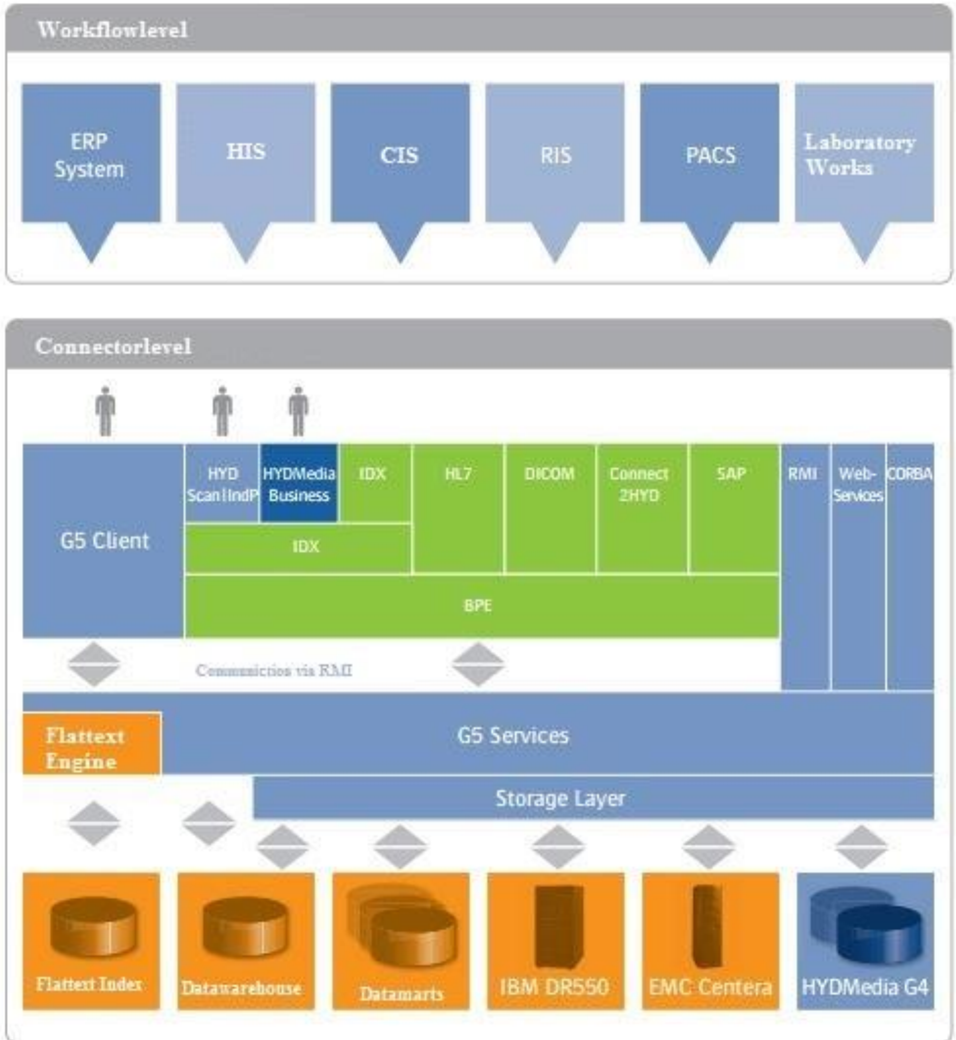

Figure 2: Enterprise Content Management HYDMedia G5 by Agfa Healthcare (Agfa-Gevaert Group 2016)

4.3 Accessing Data

It has been shown that different data sources fulfil specific purposes within the department processing it. To enable BI, Data Support and (model-based) Decision Support the DW catering to various data marts is key (Hummeltenberg 2014). The functions of a DW can be described as On-line Analytical Processing (OLAP), reporting, DM and client tools, each determined by the query and analysis components that get facilitated (Eltabakh 2012, p. 5). In order to accomplish BI's purpose of improving managerial and operational processes (Williams, Williams 2003) it is necessary to access and use the information stored in the hospitals DBMS. Some authors (Olszak, Ziemba 2006, p. 55; Lloyd 2011, p. 3) define ETL, the Data warehouse, OLAP and DM as the four key components of a BI system.

The DW itself and ETL can be seen as one unit to access single databases within the storage structure through three different kinds of application (Jing-song et al. 2011, p. 152):

1. (Traditional) Querying

The rather traditional method of using a query language extracts information from data storage. A simple query, for example, is requesting a database to display all male patients' names. Queries can be used to produce elementary reports. Aggregating information and visualizing data are useful and rather simple exercises to produce an informational overview. It is relatively simple to generate descriptive summaries within a single database by making a particular request. The Structured Query Language (SQL) is a widely used code for accessing databases and therefore querying not necessarily requires a structured DW but can be practiced on a wide range of repositories.

2. OLAP

OLAP is complementary to the DW and has the ability to not only answer the questions of "who?" and "what?" but of the "why?" and "what if?" (OLAP Council 1997). While multidimensional OLAP software reads from data cubes, relational models use a star scheme, both variations function as a decision support system granting ad-hoc data access (Eltabakh 2012, p. 24). OLAP not only aggregates knowledge but processes calculations such as percentage of total or empty beds in near real-time.

3. Data Mining

Finally DM within this thesis is defined as those methods and tools applied to generate new knowledge by using specialized algorithms to discover unseen patterns. These techniques are often associated with the terms artificial intelligence (AI) or machine learning. The fact that some authors (Alkahrouf 2005) do include OLAP into their definition of DM shows that boundaries between different tools and techniques in this sector are not necessarily well defined. Chapter 5 gives more specifications on DM and describes its applications within the healthcare sector.

4.4 Used Insight

Either if it is through mere aggregation, within OLAP or through a DM application, a hospital can derive an array of reports, functions and predictions by analysing its data. Chmielarz et al. (n.d., p. 19) show the development of Information Systems has been driven by three main principles:

- The architecture of logical systems becoming increasingly more complex
- Information systems functions suiting the individual demands of the company and its users
- Expanding spatial network infrastructure

The *Integration and Convergence Theory of Information System Development* describes the merging of functional elements to create a greater informational entity and homogenization of construction features as mayor tendencies (Chmielarz et al. n.d., p. 20). The theory explains the historical development to corporate level platforms combining Business Intelligence Systems with the ERP, including SCM and Customer Relations Management (CRM). Similar to this Serova (2012, p. 117) argues that the development of Enterprise Information Systems (EIS) has lately been driven by five tendencies:

- To automate internal processes and the relationship with stakeholders now being the central role of an ERP system
- With companies activities gaining more public interest and the use of web-based technologies ERP is bound to become more transparent
- Multilevel applications for more user-friendly options are replacing closed monolithic systems
- Enterprises of different kinds and sizes adapted informational techniques and implemented an EIS
- Deepened functionality with the drive to automate all business processes

Raphael (2014, p. 21) classifies Management Support Systems (MSS) into categories based on the information and level of support offered.

Management Information Systems (MIS) allow a detailed and multidimensional view into operative data. Within a hospital these systems are often equivalent to the CIS due to the fact that the lower management often is executed by doctors and nurses with little managerial training (Raphael 2014, p. 22).

Decision Support Systems (DSS) are modelled to support the decision making within healthcare on two stages. On a lower level patient management, treatment plans as much as the finance and inventory management is dealt with while the higher level generates supportive information with a more competitive heading (Rajalakshmi et al. 2011, p. 42). “Real” DSS consist of a Dialogue System, Model Repository, Method Repository, Database and a Report base and are scarcely found in this structure within hospitals but substituted by calculation tables such as in Microsoft’s Excel (Raphael 2014, p. 24).

Chmielarz (n.d., p. 209) reasons that information won through the use of DSS maybe further analysed within an Executive Information Systems (EIS). These are characterized by advanced user interfaces offering a choice of current relevant information to the management (Raphael 2014, p. 25). EIS are commonly based on the use of OLAP and applied in the early phases of planning or decision processes (Gluchowski et al. 2008, p. 76). They produce information that gets utilized into a balanced scorecard or function as an early alarm system by sending exception reports or using traffic-light-coding for an easy access overview for higher level managers (Raphael 2014, p. 26). EIS enable ad-hoc queries and the possibility for drill-down reports. In a hospital this means that from aggregated DRG reports the data can be accessed and analysed down to a single patient (Raphael 2014, p. 25).

Executive Support Systems (ESS) and EIS have been used interchangeably in the literature with some authors crediting ESS a higher complexity (Cano Giner et al. 2009, p. 374). According to Raphael (2014, p. 27) ESS are a mixture of DSS and EIS, with a particular focus on presenting and communicating information. Raphael (2014, p. 27) further argues that both systems are generally not well established within hospitals and therefore ESS do not get utilized to their full potential either, despite the fact that they may hold great value due to the complexity of processed data. Ikart (2005, p. 78) also describes reasons why EIS and Enterprise Information Systems - as in any MIS - do get underutilized even though they were spread across different levels of management under the label of BI or Balanced Scorecard.

Whymark (1991) relates the management level of a company to the according information type that is needed in Figure 3.

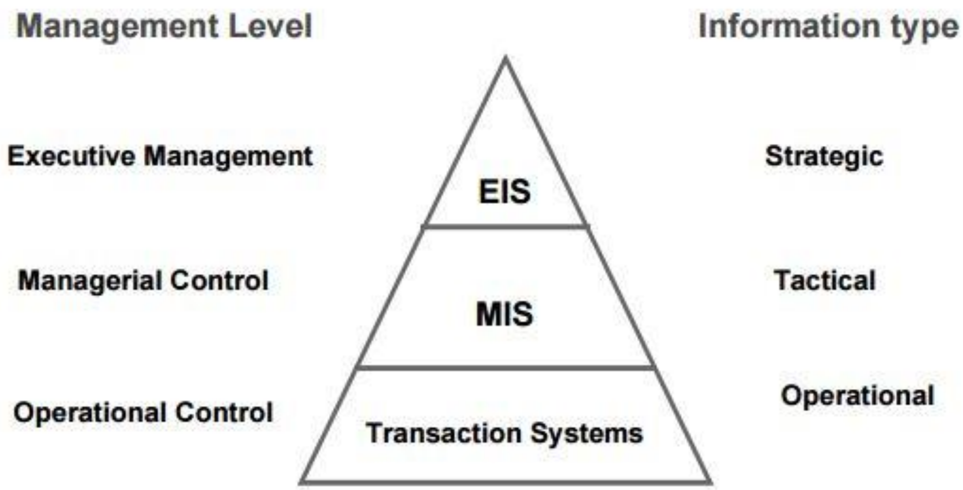

Figure 3: Management Level and Information Need. (Source: Whymark)

Transposed onto a hospital the basic transactions are the medical services provided. Within a clinic the operational control lies in the competence of the doctors and nurses, who receive and store most of their information from and within the CIS. On the managerial level the HIS is located with all of its subsets. Specific programs for controlling, logistics and HR serve the particular needs within these departments. These modules in their entity resemble a hospital's ERP for the middle management's tactical planning. In order to produce information for strategic decision making of the executive management an EIS is necessary. According support systems such as dashboards or score cards solemnly get utilized within hospitals, while their potential to optimize processes and planning to increase the economic performance is undisputed.

5. Data Mining as a Hospital Management Tool

This chapter's first part highlights the difficulties in distinguishing explorative analysis as in KDD and BI. The process of KDD and DM as its core instrument are elucidated. The second part reviews DM applications in healthcare while the third part sets the criteria used to identify studies relevant to this thesis, so that practical DM examples applicable as hospital management tools were found.

5.1 Knowledge Discovery in Databases & Cross-Industry Process for Data Mining

Analysis of data is historically connected with the development of IT and data storage techniques. Durairaj, Ranjani (2013, p. 30) argue that DM systems started in the 1960s and earlier. Due to the technical processes and developments some authors (Serova 2012, p. 116; Durairaj, Ranjani 2013, p. 31) speak of a "new generation of Enterprise Information systems" since the 2000s. With an increasing complexity of multi-dimensional storage solutions as in data warehousing and OLAP the possibilities in DM also increased.

KDD has the potential to find correlations that are not obvious and go beyond what relationships analysts may anticipate (DeGruy 2000, p. 61). The process of KDD is visualized in Figure 4:

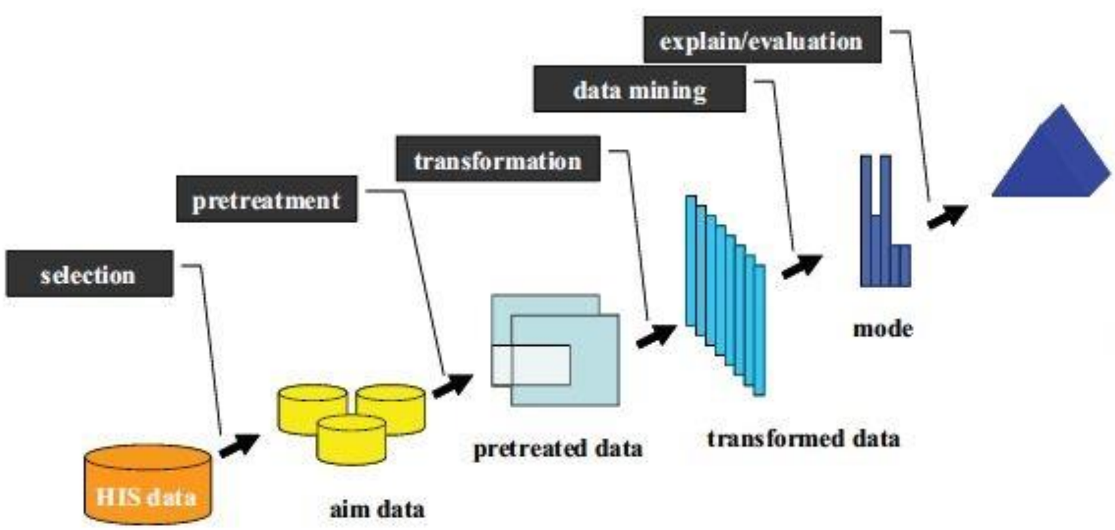

Figure 4: Knowledge Discovery in Databases (Source: Jing-song et al. 2011, p. 150)

Within this process DM is the central task of discovering such relations and for this purposes facilitates an wide array of techniques, including classification, clustering and association rules (Jing-song et al. 2011, p. 147), statistical methods, AI or machine-learning algorithms (DeGruy 2000, p. 61). KDD tools may independently determine which relationships are of significant importance but still need a domain expert to judge whether or not their business related implications are relevant to the company (Fayyad et. al. 1996, p. 42). Given the wide variety of technologies in the field it may become very hard to distinguish between BI, business analytics and "real data mining" as can be seen in an online discussion between professionals and data scientist (Quora, online). This thesis refrains to the three types of data use laid out in Chapter 4 and differentiates DM from OLAP and traditional querying as one set of tools used to generate BI. For making a distinction from the access and simple processing of data as seen in Querying and OLAP the standard Cross-Industry Process for DM (CRISP-DM) is best used to illustrate the required stages for DM. Due to the congruence in the process the steps used within the CRISP-DM are expendable onto different programs and businesses (Alapont et al. 2005, p. 2). Between 1996 and 1999 CRISP-DM was developed by three DM "veterans", Daimler-Benz, SPSS and NCR, who wanted the standard to be neutral towards industries and the specific tools used to be focused on business issues and technical analysis (Chapman et al. 2000, p. 1). Table 2 visualizes CRISP-DM's six cyclic stages including the generic task within each stage and the according outputs:

Business Understanding	Data Understanding	Data Preparation	Modeling	Evaluation	Deployment
Determine Business Objectives *Background* *Business Objectives* *Business Success Criteria* **Assess Situation** *Inventory of Resources* *Requirements, Assumptions, and Constraints* *Risks and Contingencies* *Terminology* *Costs and Benefits* **Determine Data Mining Goals** *Data Mining Goals* *Data Mining Success Criteria* **Produce Project Plan** *Project Plan* *Initial Assessment of Tools and Techniques*	**Collect Initial Data** *Initial Data Collection Report* **Describe Data** *Data Description Report* **Explore Data** *Data Exploration Report* **Verify Data Quality** *Data Quality Report*	**Select Data** *Rationale for Inclusion/ Exclusion* **Clean Data** *Data Cleaning Report* **Construct Data** *Derived Attributes* *Generated Records* **Integrate Data** *Merged Data* **Format Data** *Reformatted Data* *Dataset* *Dataset Description*	**Select Modeling Techniques** *Modeling Technique* *Modeling Assumptions* **Generate Test Design** *Test Design* **Build Model** *Parameter Settings* *Models* *Model Descriptions* **Assess Model** *Model Assessment* *Revised Parameter Settings*	**Evaluate Results** *Assessment of Data Mining Results w.r.t. Business Success Criteria* *Approved Models* **Review Process** *Review of Process* **Determine Next Steps** *List of Possible Actions* *Decision*	**Plan Deployment** *Deployment Plan* **Plan Monitoring and Maintenance** *Monitoring and Maintenance Plan* **Produce Final Report** *Final Report* *Final Presentation* **Review Project** *Experience Documentation*

Table 2: Generic Tasks (bold) and Outputs (italic) of the CRISP-DM Reference Model (Source: Chapman et al. 2000, p. 12)

1. Business Understanding

Within this first step the company's current situation is assessed in order to determine the goals to be achieved from DM. KDD does not offer "the silver bullet" to all of a company's problems (DeGruy 2000, p. 61) and therefore the sought after informational outcome should be chosen very carefully. Jing-Song at. Al. (2011, p. 161) do not include the goal-identification into their five step process for DM but point out its importance as with all software projects more complex data requires additional planning. An inventory of resources and requirements is established while assessing the situation with possible costs and benefits. Furthermore the success criteria for the business and the DM itself as well as a project plan get developed before moving on to the second stage.

2. Data Understanding

From the goals that were established the project group is now ready to determine the data requirements. The initial data collection and assessment's aim is to describe the data available and verify its quality. Exploration reports and visualizing categorical variables can be helpful towards the end of this phase (Olson, Delen 2008, p. 12). Basically the location, availability and quality of the resource data gets assessed as these factors are of major importance for the whole DM project (Jing-song et al. 2011, p. 151).

3. Data Preparation

This stage includes the steps of selecting, cleaning and integrating data into a format that is constructed to enable the modelling of a mining method. Especially the cleaning of data has proven a nontrivial and tedious task (Jing-song et al. 2011, p. 145) taking up a great share of the whole projects time particularly if data is very noisy. A meaningful interpretation of the data is only possible with this step being commenced properly and high quality data in terms of low noise and great homogeny achieved (Parvez et al. 2015, p. 46). During data preparation a more into depth data exploration may also be applied, providing the renewed opportunity to discover patterns based on business understanding attained in the first step (Olson, Delen 2008, p. 10).

4. Model Building

Once the data is prepared for mining a specific modelling technique is chosen based on business understanding gained in the first stage. All assumptions made, e.g. the attributes of a class must be numeric or attributes have uniform distributions, are recorded into an assumption catalogue. Then an actual model gets build upon these assumptions. For testing this model data is usually separated into a train and a test set (Chapman et al. 2000, p. 24). While the model gets taught on the train set, its validity and reliability gets estimated performing on the separate test set. Olson, Delen (2008, p. 15) point out that for initial analysis visualization tools are most helpful to plot data and establish relations. Cluster analysis can be used to help identifying which variables are well correspond; while more detailed models can be applied as the project evolves (Olson, Delen 2008, p. 15). Commonly the DM engineers apply a technique multiple times or try a string of techniques to finally compare the results according to the initially set success criteria (Chapman et al. 2000, p. 25). For a more accurate disease prediction, for example, a combination of DM techniques has to be utilized (Parvez et al. 2015, p. 46).

5. Testing and Evaluation

Factors like validity and accuracy of the model get evaluated during the model building phase (Chapman et al. 2000, p. 26). In the fifth stage results from training and testing of the build models will be evaluated according to the success criteria established. As an outcome other informational needs will be identified, regularly reverting to prior phases of the whole cycle (Olson, Delen 2008).

6. Deployment

DM may not only be used to discover unseen patterns and non-assumed relations but is used to verify previously made hypotheses. While valid models can be expended onto different business operations for many purposes it is also important to revise those models according to changes in the operating conditions (Olson, Delen 2008).

Contributing the fact that DM projects may need a long time and require regular revision CRISP-DM is a cyclic work frame.

5.2 Data Mining Techniques

A recent review on DM techniques in healthcare is by Parvez et al. (2015, 39ff), who categorize all tools into four classes, explain their functionality and outline their common utilization:

- Classification

Unarguably one of the most common DM methods in healthcare, classification determines different target classes within the sampled data. A class prediction is made for each case, e.g. the analysis of disease patterns in order to define a patient's risk factors. K-Nearest Neighbour is an easy and efficient classification algorithm that gets trained and is memory based (Tomar, Agarwal 2013, p. 244). Support vector machines use Kernel functions, Neural networks so-called neurons or nods and Bayesian Methods are based on the Bayes Theorem (Parvez et al. 2015, 40f). Finally the Decision Tree is a flowchart like, non-parametric method that is also included in the classification category. Due to its high sensibility to the training data, irreverent attributes and noise decision trees show a low performance on complex data sets (Parvez et al. 2015, p. 40) but used on breast cancer prediction were also "found to be best predicator with greater amount of accuracy as compared to other techniques" (Kundu, Pallavi 2015, p. 904).

- Regression

The mainly mathematical regression instruments are well known from the field of statistics. While linear regression is applied to identify the correlation of two numeric variables, non-linear or logistic regression can be applied to categorical data and predict the probability of an occurrence (Parvez et al. 2015, p. 43). In the medical field regressions are commonly used to predict a disease or a patient's chance of survival (Tomar, Agarwal 2013, p. 250).

- Clustering

Clustering uses similarity measures to group the database into different subsets, partititional, hierarchical and density based clustering are the algorithms used in healthcare (Tomar, Agarwal 2013, 251ff). Other than regression, clustering only observes the independent variable. As clustering has no predefined classes it is best used for exploratory studies dealing with huge data that is little known about (Parvez et al. 2015, p. 43). Such may be dealing with the human genome or biological taxonomies having the identical purpose of establishing categories and assigning the cases accordingly.

- Associations

The first association rule algorithm was developed in 1991 but due to its technical difficulties did not receive much attention until IBM devised the *Apriori* Algorithm (Parvez et al. 2015, p. 45). Association rules are now an essential DM technique in order to identify frequent patterns and significant relations among data sets within a repository. In the medical field it is often used to detect relations between a disease, its symptoms and a patient's health status or utilized by healthcare insurance companies to detect fraud and abuse (Tomar, Agarwal 2013, 255f).

These are the four categories of DM as defined in Parvez et al.'s (2015, p. 39) article. Reading into the literature an extensive amount of different algorithms for all the categories and their sub classes can be found.

Durairaj & Ranjani (2013, p. 32) differentiate DM techniques into just two classes. Being either *Predictive*, including prediction, classification, regression and time series analysis models or *Descriptive* methods, which use models for clustering, summarization, association rules and sequence discovery. For the descriptive methods the discovery of correlations in a data string and detecting exemptions and nonconformities are the activities DM gets most frequently associated with, while predictive analyses find their most beneficial applications through the facilitation of user interfaces (Olszak, Ziemba 2006, p. 50).

Moreover DM techniques can be distinguished as unsupervised learning, only monitoring the independent variable or supervised learning which observes both, the independent and the dependent variable (Parvez et al. 2015, p. 43).

To decide on the appropriate learning techniques and the attributes used in the model are key factors for the project established in the first three steps of the CRISP-DM. Unrelated to these decisions and whether the data gets accessed from a warehouse or through simple querying it is likely that a utility program is used to convert the mining data into the required format (Jing-song et al. 2011, p. 145).

DeGruy (2000, p. 67) points out that KDD does not substitute for business analysists but may make their work more efficient and thus can be a competitive advantage when suitably applied. Furthermore clinicians may feel uneasy at the thought of data scientists elaborating models to mine medical data without having the according knowledge of diseases, their effects, treatment and common procedures. On the other hand data miners may interpret their lack of understanding the medical background as a benefit because they therefore would not be biased by expecting certain results. As Baca-Garcia et al. (2006, p. 1128) put it: “The machine learning multivariate methods are totally blind and unbiased.”

Compared to SQL and OLAP, DM goes a step beyond accessing and reading data by facilitating the different algorithms to already analyse the data and consequently it is the most sophisticated approach to use data for generating BI. Through the high workload of defining objectives and data preparation a mining project usually is expensive. This is especially the case when DM is applied to a hospital as many meetings have to be held with a variety of specialists from different departments (Alapont et al. 2005, p. 1). Biron et al. (2014, 195f) give an example for a relatively low cost project with a clinic’s IT department developing a full-text search tool from an open source program in collaboration with a private software company. The whole venture is calculated at around €15,000 and a length of six months with the biggest cost factors being the compensation for the clinic’s developer finishing data integration (€7000) and wages for two trainees (€5000), while the remaining €3000 were spend on purchasing a new server (Biron et al. 2014, p. 195).

Setting up the IT structure with interfaces to all subsystems and a central DW is the most basic task for a DM project and depending on the hospital's presuppositions could also prove the most labour- and cost intensive one. For extraction of knowledge from this infrastructure the access through DM takes the highest amount of determination and technical knowledge to be applied appropriately but may also allow the greatest insight as it is specifically targeted to identify unseen patterns. Hospital managers could utilize the mined knowledge in order to provide a better planning and ensure the hospital's development (Jing-song et al. 2011, p. 155). Intending to understand the potential of DM applications specifically used as a management tool, it is advisable to search for practical examples described in the literature.

The following Chapter defines how relevant studies were assessed given the difficulties in technical differentiation of some terms and the high volume of studies as they are related to aspects of healthcare, IT and management.

5.3 Data Mining in Healthcare and Categories for Reviewing Relevant Literature

A wide variety of studies regarding the benefits of DM in the healthcare sector exists. The 2015 literature review by Parvez et al. (2015) examines a total of 47 studies according to the mining techniques used and their intended purpose. The papers included in this review were summarized to gain an overview of possible DM applications in healthcare and identify focal points of the current research. Twenty-eight of the studies mentioned in Parvez et al. (2015) used classification algorithms of which twenty-four could be filed under the topics treatment effectiveness or the general healthcare management. The models build are often used to assess and predict risks. Clinical DSS for practitioners are derived as much as systems for the discovery of diseases or their early detection, ranging from various sorts of cancer, over heart and lung diseases to dengue fever. Two of the classification studies regard the CRM, as in predicting treatment costs and the services chosen by patients with chronic conditions. One concerns detection of fraud and abuse by users and insurance companies. However, only one paper is targeted to the hospital management directly. Govaert et al. (2015) show that surgical auditing may improve a hospital's economic performance on surgeries if only offered for high-risk procedures. Nonetheless this paper is a systematic review and it remains unclear why it is included into Parvez et al.'s paper as using a classification algorithm. Furthermore four studies using regression models are mentioned. Two of them effecting the patient's treatment respectively healthcare management and two regard the CRM or hospital management directly. Xie et al. (2015) developed a regression model accurately predicting days of hospitalization within a population and its subgroups, based on extensive data from insurance claims. With data from only one HIS Alanpont et al. (2005) showed that regression models may have a higher accuracy in predicting average admissions to a hospital's Emergency department then the traditional statistical methods. The nine studies utilizing Clustering Algorithms seem more diversified in their purposes. Four of them focus on treatment effectiveness, all addressing microarray, which is highly specialized genetic data. Two papers deal with fraud and abuse detection, and another two are aimed at the cost prediction and identification of homogenous patient subgroups, therefore concerning the CRM and general healthcare management. One of the Clustering applications may be interpreted as a hospital management tool with Belciug (2009) using a model to group patients according to the length of their anticipated stay in order to improve the services provided. All of the six studies using association as DM technique finally can be aligned to treatment effectiveness, healthcare management and CRM. Associations studied are between the diagnosis and treatment, drugs and adverse effects and a clinical DSS based on an association rule shall help to detect non-communicable diseases in remotely located communities.

Using this and other reviews it is fair to conclude that the biggest portion of health related DM applications examines the data in search for an optimized treatment or the better understanding of risk factors. It is important to acknowledge that all information gained to understand a disease and decisive factors in its treatment may benefit patients and therefore holds implications for the clinical practice. Same accounts for studies focusing on the improved prediction of a disease's prevalence and the services patients use as they may help the hospital to estimate future demands.

Baca-Garcia et al. (2006) show that DM methods may have higher accuracy determining whether a suicide attempter has to be hospitalized or not then the traditional statistical methods applied by psychiatrists. The authors were able to build a classification model from six feature selection methods that could classify 99% of the cases from a public general hospital in Spain correctly. Yet the article does not claim that a DM approach to this classification problem is more valid then the clinicians' assessment but merely that therapists and psychiatric researches may need to recognize these techniques and work with DM engineers. All the same this is just one example on how hospitals may utilize studies generically commenced to improve the operational proceedings on all managerial levels. As the middle and upper management have to ensure that available knowledge and processes positively influencing the operational functioning of the hospital get facilitated, chief executives shall also pay attention to studies identifying fall risks or implementing a model for better triaging of patients. It is the executives management's strategic decision to eventually implement a system that improves the operational functioning, like a Cancer Centre in Lyon developing their own text mining system to improve and streamline nurses and practitioners access to medical information (Biron et al. 2014, p. 197). Valuable knowledge for the tactical and strategic planning directly may be found in applications for a more accurate prediction of patient's lengths of stays, services demanded within a clinical pathway or the improved use of resources. To serve as practical examples regarding this thesis' topic the focus is on studies that

- **(a)** test DM applications that got developed to support a hospital's tactical or strategic management directly and
- **(b)** were primarily run on one hospital's or clinic association's HIS.

Although they may hold very real value for the institution, papers focusing on an improved treatment effectiveness which may lead to a better operational functioning of the hospital did not get considered. Also studies that deal with the set-up of data warehousing solutions or management of medical records and their general business significance were excluded. The benefits and importance of DBMS within the hospital sectors already were well established and this review's purpose is to find possible practical DM applications on HIS for the strategic planning. Even though DM applications on HIS have been described as early as 1997 - Prather et al. identified factors contributing to preterm birth - only studies published from 2000 on were included as this is when the CRIDP-DM got established as standard procedure for data mining. A number of such submissions can be found and is listed according to the source database in the Tables included in Annex B.

6. Findings & Discussion

This Chapter presents conclusions drawn from the literature reviews and the interviews conducted. The first part examines the possibilities and examples for analysing the HIS structure within a country and grants an insight into available Austrian data. The second part focuses on the potential use of DM as a hospital management tool.

6.1 Data Processing in Austrian Hospital Information Systems

Intending to assess the IT structure within Austrian hospitals a search for studies regarding this topic or any that were about the assessment of HIS on a national level was conducted. As just single components of the system may be examined difficulties originated from the complexity of the whole HIS. Furthermore this thesis focused on the extraction of business related knowledge from HIS in order to create BI, a rather specified application considering the main medical and administrative tasks of these systems.

The following chapters examine some contemporary reports on IT usage in (German) hospitals, give an outline of other related projects in Austria and gauge publically available information from an organisation certifying hospitals. Finally four studies evaluating HIS on a national level were described and two well documented Austrian HIS examples, the Tyrolian State hospitals and their Styrian equivalent were further elaborated upon.

6.1.1 Contemporary IT Report on Hospitals

Regular reports on the IT use in the public German healthcare sector have been published by Hübner and colleagues since 2004. The reports have been developed to understand the process and status of IT Systems in hospitals by analysing the extent to which clinical and administrative functions are facilitated within IT. Two of these reports (Hübner et al. 2010; Hübner et al. 2015) with a specific focus on nursing also include data from Austria.

The earlier study (Hübner et al. 2010, p. 6) shows a comparably high penetration of IT functions with 91,1% of the responding 45 Austrian hospitals having an IT-based patient management and staff-scheduling system. Also widely introduced were PACS (82,2%) and radiology information systems (77,8%). 66,7% of the hospitals reported to have a nursing documentation system, 62,2% clinical workstations. Finally the first comparative study finds Austria seemingly ahead of Germany regarding the facilitation of IT in public hospitals. With financial sources ruled out as a crucial factor for development the authors could not conclude either if smaller countries have an advantage when introducing innovations into the practice or that differences are caused by other reasons such as legal constraints, alterations in the technical infrastructure or the systems itself, like a considerably lower length of stay in Austrian patients hospitalized (Hübner et al. 2010, p. 11). That legal regulations play an important role for the widespread facilitation of IT based nursing documentation was also suspected by Ammenwerth (Ammenwerth; Schmied 6/3/2016, 46ff).

The tendency of Austrian hospitals better integrating IT functionalities then the German institutions generally prevailed in Hübner et al.'s later report, especially for the electronic nursing documentation which was utilized more than twice as often in Austria (67,1%) than in Germany (31,2%) (Hübner et al. 2015, p. 139). 70 Austrian hospitals answered the authors questionnaire delivering contemporary knowledge on which functions were covered by IT. The transmission for requests of performance and feedback of the according results from five diagnostic departments were evaluated. Laboratories were most likely (63,2%) to be found fully connected by IT, followed by the radiology (50% without, 45,6% including images) departments; other diagnostics such as echography (30,9%) and 19.1% of the electrophysiological diagnostics were fully integrated (Hübner et al. 2015, 85ff). Notably these results only include the cases with integration across all wards, not those that already started a pilot project or were planning to do so. Percentages of hospitals that do not plan particular IT projects were low, e.g. only about 6% did not plan to develop an IT strategy for either the radiology department including images or the sonographies, while big proportions of the hospitals (11,7 – 35,5%) made no comments on these questions. Concerning three functions for decision support 21,5% of the Austrian hospitals fully integrated medical guidelines and clinical pathways (CP), 20% offered alarm functions, such as automated notification for conspicuous lab results and 15,4% clinical reminders (Hübner et al. 2015, 88f). Also notable was a relatively high number of hospitals not preparing for expansion of the IT functions with 30,8% for Guidelines and CP, 27,7% for alarms and 40,5% of the hospitals not planning to develop clinical reminders. An computer based Clinical Incident Reporting System (CIRS) was available on all wards in 38,5% of the responding hospitals, a system for tracking the patients in 32,3%, while medication only got traced in 6,3% of the clinics (Hübner et al. 2015, 90ff). Material supply, the pharmacy and catering for patients were fully integrated IT functions in about 65 to 70% of the hospitals, with only 3,2% not planning to swap to a computer-based system for supplies and pharmacy (Hübner et al. 2015, 93f). Commonly the planning of nursing shifts (82,3%) the management of in-house (77,4%) and outpatients (51,6%) was IT-based, which was only the truth for 12,9% of the hospitals when it came to communicating to other healthcare providers (Hübner et al. 2015, p. 95). Of particular importance for this thesis is the fact that 52,5% of the Austrian hospitals stated to have a secondary use of the data for economic planning, treatment analysis or research and education. In most of these cases (93,5%) the secondary use was annotated as for improving patient safety or quality management, all percentages for the purpose of the secondary data use comparing Austria and Germany can be seen in Figure 5:

Figure 5: Types of Secondary Data Analysis in Country Comparison (Source: Hübner et al. 2015, p. 144)

Most notably 67,7% of the secondary data analysis cases included the purpose of management control and data warehousing, which fulfils the definition of generating BI. Accordingly over half of the Austrian Hospitals used clinical files for further analysis with close to 70% of these efforts also aiming at BI. Still uncertainties remain about the nature and quality of the analysis, whether they were retrospective or prognostic. It also persists to be unclear if the generated business knowledge was simply aggregated in reports and visualized as key indicators and graphs or facilitated in some kind of managerial dashboard with alarming functions. Through the fact that in Germany secondary data use was more common (71,4% as opposed to 52,5%) but the proliferation of purposes was smaller (54,4% to 63,1%) the authors concluded that when secondary analysis of data is commenced in Austria the analysis then is more comprehensive and into depth (Hübner et al. 2015, p. 145).

6.1.2 Related Projects within Austria

Even though it is fair to assume that Austria's structure of HIS and the use of BI is fairly developed no studies particularly assessed public hospitals IT capabilities in terms of extracting business relevant information. Neither was it possible to determine if hospitals develop their own structure for managerial DSS or rather use commercial tools. Management tools such as Balanced Scorecards or Dashboards are underutilized within the hospital environment (Ikart 2005, p. 78; Raphael 2014, p. 24) and through the lack of empiric studies it was not possible to conclude why these, in other industries well established management tools, do not get fully exploited.

In Austria an initiative consisting of professionals and experts from public and private institutions was founded in order to define a road map to a modern Austrian Information and Communication Strategy for healthcare. This E-Health Initiative (EHI) is organized into seven different working groups and the gradual introduction of an Austrian electric patient file, so-called ELGA, may be seen as the central goal to achieve and has been strongly promoted by this initiative. One of the key aspects for the EHI is the exchange and interoperability of data between different healthcare subsets and providers generally and the workgroup on structure related information systems includes DM and meta-analysis specifically into their highly prioritized fields of interest (Österreichische e-Health Initiative. Jänner 2007, p. 26). Still the EHI's 2007 positioning paper does not elaborate on useful DM applications and specific themes for implementation, neither does it give indications on the status of Austria's HIS. The report highlights the importance and potential paybacks from centralized data storage and an enhanced knowledge transfer, pointing out that standards in format would be highly advantageous. That medical standards like HL7 or DICOM get fully utilized got confirmed by all of the interviewees that deem these formats to be very helpful, while their maintenance may require some effort as they sometimes “speak different dialects” (Kocever; Schmied 6/6/2016, 52f).

With the *HIS Monitor* “[a]n approach to assess the quality of information processing in hospitals” (Ammenwerth et al. 2007) has been undertaken in Tyrol. Based on reviewed literature and the author’s own experience a two axis matrix has been developed. One axis reflected a patient’s pathway from admission procedures, treatment selection and documentation to discharge or transfer while the second axis described six criteria of informational quality: availability, correctness & completeness, readability, usability, compliance with legal requirements and the time needed for processing of information (Ammenwerth et al. 2007, p. 218). So within this matrix the authors could formulate a question that assesses any of the six quality criteria for each of the recognized patient steps and then identified the most relevant of these questions and refined their formulation through various pre-tests. Feasibility was tested in a surgical and an internal medicine department with the HIS Monitor including 107 + 10 questions of which 81 had to be answered by physicians, 77 by nurses and 20 by administrative staff. Thus the author’s acknowledged certain possible cofounders the tool was evaluated to be helpful with assessing the user’s perception of an HIS. Furthermore the subjective assessment may be more important than merely quantitative surveys as there are examples of staff rejecting new systems that were not user-friendly and that the best system may be impracticable if it was badly perceived by staff (Ammenwerth et al. 2007, p. 223). Rather than identifying the HIS’s technical structure and therefore not analysing reasons for good or bad performance the HIS Monitor gives indications on how the users anticipate a certain system. In 2006 the Tiroler Landeskrankenanstalten decided to change from a paper-based to a digital nursing documentation system and a subset of the HIS Monitor with 41 + 10 questions to nurses was used in order to assess how they perceived the changes in documentation quality with the change of the system (Ammenwerth et al. 2011). The IT based nursing information system clearly improved processing of information; even though one third of the nurses felt that this system was more time consuming, while half said it would save time (Ammenwerth et al. 2011, p. 30). The study also refers to a few other researches with similar findings on the implementation of a new nursing documentation system, while the HIS Monitor seems to be the only existing survey that covers different processes along the patient’s pathway. With further research the authors argue this tool may be refined and made more useable through shortened questionnaires to finally be used to evaluate the different stages a healthcare institution goes through in adopting IT (Ammenwerth et al. 2011, p. 34). Despite some keen interest from hospitals, with requests coming as far as from Sweden, the project was discontinued as it was not refunded. The lack of a screening instrument to benchmark and compare nursing information systems between hospitals therefore still remains (Ammenwerth; Schmied 6/3/2016, 35ff).

6.1.3 Cooperation for Transparency and Quality in Healthcare (KTQ)

More than 10 years ago the KTQ "was founded by public German institutions to implement a solution for quality management that includes certification for German medical facilities" (KTQ International 2016). Next to the *Technischer Überwachungsverein* which uses ISO standards now the KTQ is one of the major hospital certification agencies. KTQ focuses on six central categories in healthcare based on which the institutions do a self-assessment according to the four steps of the PDCA cycle, judging on the level of achievement and the level of pervasion for each category and step. Within a German clinic's report the topics on assessing the category 'Information and Communication' as in the latest KTQ Manual Version 5.0 can be found:

Subcategory	Criteria
4.1 Handling of patient data	4.1.1 Regulations on distribution, documentation and archiving of patient data
	4.1.2 Documentation of patient data
	4.1.3 Accessibility of patient data
4.2 Transmission of information	4.2.1 Transmission of information between different departments
	4.2.2 Transmission to and from central information offices

4.2 Transmission of information	4.2.3 Public information dispersion
	4.2.4 Consideration of data security
4.3. Usage of Information Technology	4.3.1 Structure and usage of IT

Table 3: Assessing *Information and Communication* as in the KTQ Manual (Source: Kooperation für Transparenz und Qualität im Gesundheitswesen 2015)

To each of these criteria questions enabling the regarding healthcare institution to assess the according PDCA steps determining the level of achievement and pervasion were formulated.
On KTQ's official webpage (www.ktq.de) currently 21 Austrian Hospitals are certified with their according quality report accessible. While most of these reports read similarly they are, due to the self-reporting characteristics of KTQ, very much varying in the depth they touch on a certain aspect of the information and communication structure. Summarizing the reports, the hospitals generally seem to have dedicated IT and data safety concepts. The majority of reports mention that an in-house e-mail service or Intranet is available for internal communication, while external presentation is commonly done through webpages, reception desks or leaflets. The HIS were permanently accessible, despite maintenance intervals, sometimes back-up and emergency concepts get mentioned in case that the servers went down and usually a firewall protects the hospitals from viruses and spam mails. The importance of such protection cannot be underestimated as hospitals do fall victim to cyberattacks and "anyone considering themselves safe does not have a clue" about the threat (Kocever; Schmied 6/6/2016, 122f). All reports emphasize the fact that the handling of data and archiving the patient files is according to all relevant laws. Access was mostly regulated through dedicated user accounts and passwords, while one hospital claimed to use biometrical logins in form of fingerprints. A focal point is the EHR/ELGA, with only one report not particularly mentioning it. Most of the times the EHR was either "in the focus" of the institution, a "dedicated strategic goal" to include or "currently under integration". Only two institutions said they have fully adopted the EHR, while the partial use of it often seems to be defined as having an EHR but still using paper-based

fever charts. In most of these cases the fever charts, containing basic information about daily routines such as vital signs and medication given, do get scanned and transferred into the HIS after the patients discharge. Upon this transfer in some institutions the data does get checked for plausibility and comprehensiveness. Only one institution named the particular software used for patient documentation.

Through these reports insight into individual IT solutions could be gained as much as a general overview of those structures and the use of paper-based files. Yet, these discernments are limited to those facilities participating in KTQ`s certification schema.

6.1.4 HIS Assessment Methods on a National Level

Four papers assessing the nationwide status of the HIS structure were identified. In 2007 Velez Lapão surveyed the Portuguese HIS status using a modified Nolan maturity Matrix. Richard L. Nolan devised a theoretical model on the growth of a businesses' IT in the 1970s and his maturity index is well-recognized for providing the theoretical foundation for characterizing the maturity of an informational system (Mykityshyn 2007, p. 111). Nolan's maturity Index for IT systems serves as theoretical base for a variety of related studies and is pictured in Figure 6:

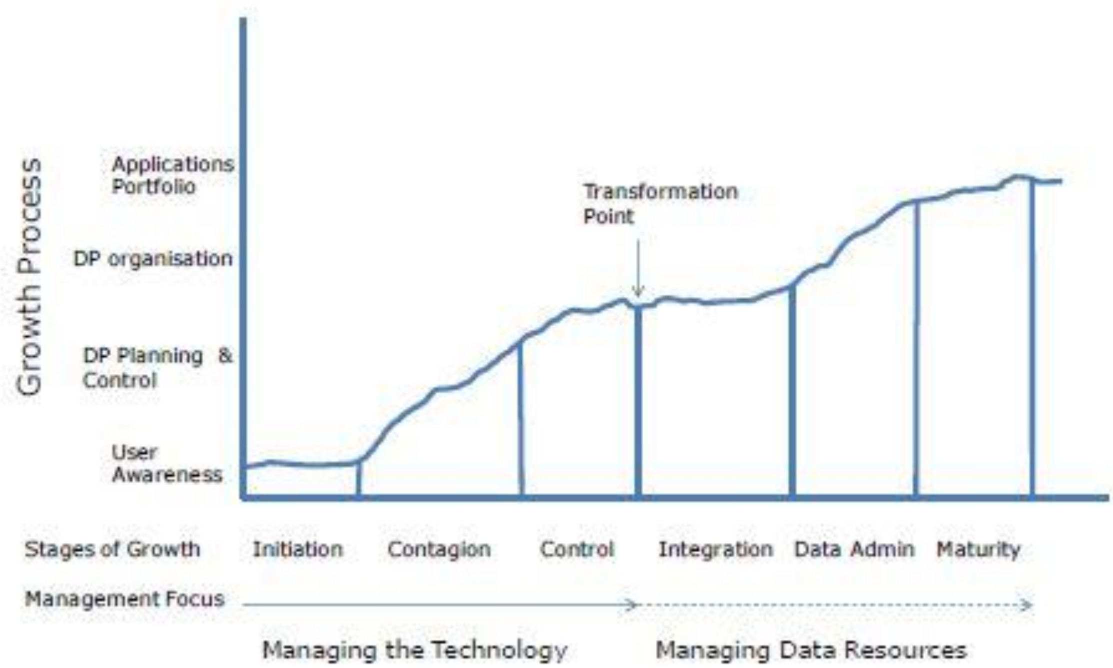

Figure 6: Nolan's 6 Stage Growth Model (Source: Galliers and Sutherland, 2003 amended from Nolan, 1979 in O'Neill 2011, p. 15)

Velez Lapão used a healthcare adapted matrix and piloted the survey with the HIS department's heads (HISDH) in 73 national public hospitals in 2005. The response rate was 41%, with most of the hospitals not responding due to the fact that either the function of CIO/HISDH was outsourced or the persona was always busy and could not find the time to reply. In fact the authors show that a lack of skilled personal seems to be one of the major barriers for a more mature HIS system in Portugal. Numbers of staff in IT departments were comparably low and only 27% of the HISDH had their position for more than five years, indicating a lack of experience (Velez Lapão 2007, p. 494).

Another study examined the status of Iranian HIS (Jahanbakhsh et al. 2014). It describes the situation in seven Iranian Hospitals as they have been assessed through teams of interviewers, which at the same time were professionals and lecturers in e-Health. Not surprisingly HIS in the developing country Iran seems to face more basic difficulties than in Europe. An essential problem lays in the fact that law requires a great share of information, e.g. surgical records, and insurance claims, to be transmitted in paper form (Jahanbakhsh et al. 2014, 273f).

In order to explore the *Health Information Technology* (HIT) *Innovativeness* a Canadian study (Paré et al. 2009) developed a measure consisting of eight dimensions in three different vectors, shown in Table 4.

Vectors	Dimensions		# of items in instrument
Functional innovativeness	D1	Administrative systems	8
	D2	Patient management systems	8
	D3	Clinical support systems	4
	D4	Clinical systems	10
Technological innovativeness	D5	Emerging technologies	13
Integration innovativeness	D6	Internal integration – administrative	1
	D7	Internal integration – clinical	1
	D8	External integration	4

Table 4: Structure and Components of the HIT Innovativeness Measure (Source: Paré et al. 2009, p. 31)

An overall of 49 items were compiled and the authors received a total of 106 questionnaires (response rate 52%) completed by the CIO of the public Canadian hospitals, sensing a "strong support for the research method" (Paré et al. 2009, p. 28). Overall the authors tested nine hypothetic factors influencing the hospital's IT Innovativeness, finding seven of them to have significant impact. First is size, as lager hospitals can allocate more resources to the acquisition and maintenance of new systems and technologies (Paré et al. 2009, p. 33). Accordingly finances are a "very important factor" to determine HIT innovativeness in all three vectors, while the CIO's IT experience and presence of an dedicated IT steering committee appeared to significantly influence the implementation of new systems but not the HIT integration (Paré et al. 2009, 33f). Availability of IT personnel also had an impact on all three layers of the innovativeness. Urban hospitals were found to have a better functional innovativeness than their rural counterparts, while this geographical factor did not seem to influence the technical or integration dimensions (Paré et al. 2009, p. 33). Interestingly network affiliation was found to have no significant impact on the IT development and the fact that a US study came to a different conclusion may be explained by a different set-up for budget allocation within the hospital networks in those countries. Furthermore an affiliation to teaching was not significantly correlated with the HIT innovativeness, except for a higher technological innovativeness the teaching hospitals did not show a higher number of implemented IT systems or a better integration of those (Paré et al. 2009, p. 35). Overall the authors conclude the participating hospitals to have a "moderate level of HIT innovativeness […] with more emphasis on administrative systems and their integration than on clinical systems and technologies in these settings" (Paré et al. 2009, p. 36).

Finally a dissertation by O'Neill (2011) undertook an heuristic assessment of the Irish hospital's HIT capabilities. The author also adapted Nolan's Maturity Matrix and Paré's methodology, deriving specific models to assess a hospital's capability within nine dimensions:

- Electronical Medical Record (EMR)
- PACS
- Enterprise Resource Planning (ERP)
- Interoperability
- Project Management
- Program and Portfolio Management
- Hit Use and Management

- Enterprise Architecture
- Quality Management

The capabilities in a hospital’s ERP, for example, were assessed by analysing how IT is used for strategic planning, the organisation’s sophistication, the extent to which the ERP got adopted, the hospital’s vision and what the motivational drivers and learning processes have been (O'Neill 2011, p. 22). HIT use and management were assessed by questions covering the organisational overview, the networking with external information, as much as the integration of managerial and healthcare technology (O'Neill 2011, p. 29). The CIO’s responding to the study had to state whether or not a particular function was implemented into the IT and if so to which extend it got used. Ratings were on a 7 point scale ranging from 1 (barely used) the computerized system gets used around 10% of the times, to 7 (extensively used) indicating a usage of at least 80% of the time. Table 5 shows the corresponding figures for the usage of an Executive Information system or BI applications respectively.

	No Plan	Plan	Installed not used	1 Barely used	2	3	4	5	6	7 Extensively used	Rating Average
Financial/Clinical dashboards (executive information systems)	22.2% (4)	11.1% (2)	0.0% (0)	0.0% (0)	5.6% (1)	5.6% (1)	**27.8% (5)**	11.1% (2)	0.0% (0)	16.7% (3)	3.06
Business intelligence applications (e.g., OLAP cubes, data mining tools)	5.9% (1)	29.4% (5)	0.0% (0)	0.0% (0)	5.9% (1)	5.9% (1)	**35.3% (6)**	5.9% (1)	5.9% (1)	5.9% (1)	3.06

Table 5: Financial/Clinical Dashboards & BI Applications in Irish Hospitals (Source: O'Neill 2011, p. 120)

Notable is a relatively high percentage (22,2%) of the hospitals did not plan to introduce financial/clinical dashboards, while only 5,9% did not plan to us BI applications. With 29,4% BI systems are also ranked highest among those systems that are planned for implementation (O'Neill 2011, p. 58). Eleven hospitals (61.1%) indicated to plan the implement of an ERP system, “such as SAP or Oracle”, one (5,6%) started implementation and six (33,3%) stated to have such a system readily available (O'Neill 2011, p. 68). The author furthermore found only 14% of the respondents used a Business Process Modelling tool and 79% had no plans to do so, arguing that the possibility of best-practice management of HIT integration and business processes “is being overlooked by the majority of hospitals” (O'Neill 2011, 74f).

Finally the dissertation included a scoring mechanism developed to calculate the HIT capability (HITCAP) for the individual sectors and organisations while enabling comparability with similar studies. A descriptive summary of all participants (n=20) HITCAP Scores can be seen in Table 6:

Vector	Dimension	Minimum HITCAP	Maximum HITCAP	Mean HITCAP	Standard Deviation
Functional	Admin Systems	0	88.8	**58.7**	23.1
	Patient Systems	48.5	100	**77.7**	13.7
	Clinical Support	50	100	**76.2**	18.5
	Clinical Systems	24.2	84.2	**49.0**	17.1
	Vector Values	49.3	83.9	**65.4**	11.4
Technological	**Technology**	0	61.5	**32.6**	16.5
Integration	Internal - Admin	0	100	**29**	40.7
	Internal – Clinical	0	100	**40**	37.8
	External	0	64.4	**26**	14.9
	Vector Values	0	76.3	**31.6**	21.4
Sample Values		26.1	74.9	**48.6**	11.8

Table 6: HITCAP Score Summary Descriptive Analysis (Source: O'Neill 2011, p. 54)

The authors came to the conclusion that generally speaking “HIT in Ireland is stuck rigidly in the Control phase” of the Nolan Growth Model (O'Neill 2011, p. 99). With the existence of valid Maturity Models for the different domains of HIT the author further suggests to make the according questionnaire available as in an online self-assessment that could serve as basis for a national register of systems and standards in use so that areas needing attention could be identified more easily (O'Neill 2011, p. 103).

6.1.5 Two Prominent Austrian Examples

The development of at least some components of an HIS within two of Austria’s major hospitals networks were documented in the literature. Namely the *Tirol Kliniken GmbH*, which at the time of the studies was still called Tiroler Landeskrankenanstalten GmbH (TILAK) and the Styrian equivalent the *Steiermärkische Krankenanstaltengesellschaft m.b.H*. (KAGes)**.**

The TILAK compromises four hospitals with an over-all of nearly 2500 beds in the state of Tyrol dealing with more than 118.000 patient admissions in 2014 (Tirol Kliniken n.d., p. 5). The above mentioned introduction of an IT-based nursing documentation and evaluation of quality changes using the HIS Monitor was part of a general program in which the TILAK defined its specific IT strategy from 2003 to 2007. Already beforehand a well-established and sophisticated IT system, with most diagnoses transposed into the company-wide CIS, a structured PACS and broad IT support for economic tasks was in place (IT-Strategiebeirat der TILAK 2002). SAP-ISH is used for the patient administration and billing while SAP-R/3 is utilized for the business administrative and economic functions (Lechleitner 2007, p. 14). Furthermore the TILAK has been one of the first customers installing the private Company *Cerner*'s software *Millennium* as a CIS (Stark; Schmied 6/1/2016, 31ff) and an advanced image management (AIM) originally based on Siemens *syngo.share* is consolidating classic RIS and PACS systems connecting all hospitals within the Tirol Kliniken. A program called *Cloverleaf* serves as an interface between CIS and AIM and may be seen as a converter of files and formats between all different subsystems, having around 3000 daily users. Figure 7 is a visualization of the connections between TILAK's around 400 single components via Cloverleaf and gives an impression of the complexity of such arrangements.

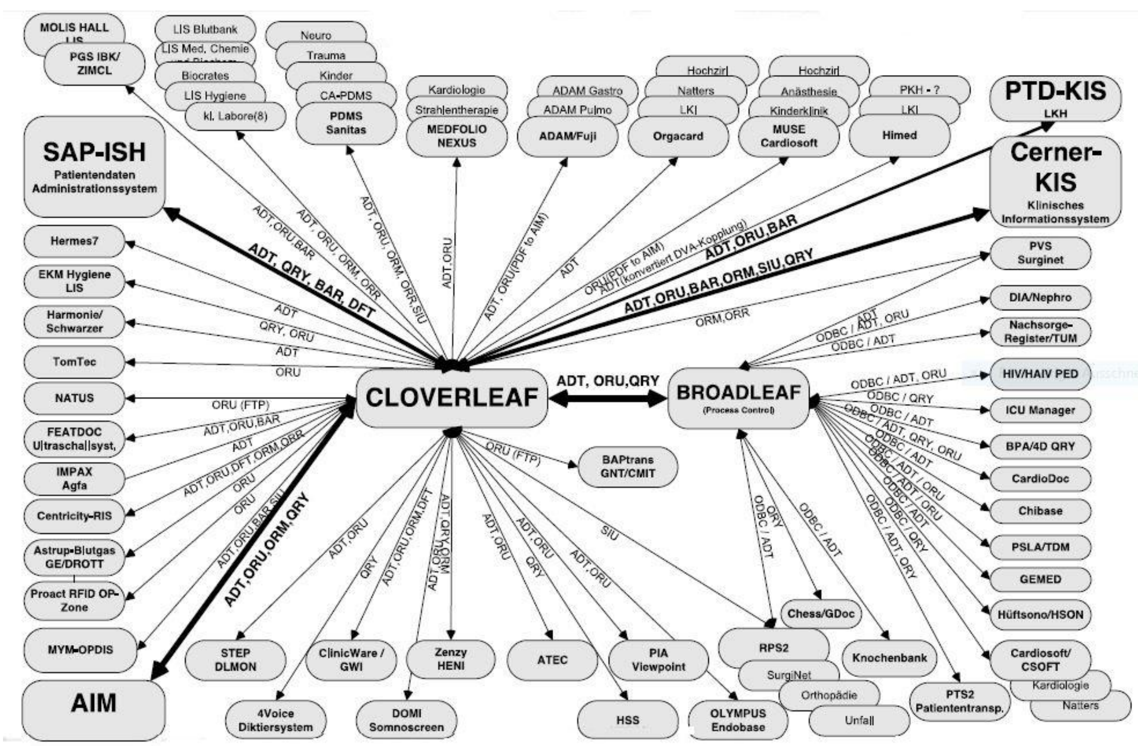

Figure 7: HL7 Streams at TILAK with Communication Server CloverLeaf as Central Information Broker (Source: Lechleitner 2007, p. 12)

The proposed IT strategy was aimed to reduce the amount of specialized IT solutions used in the medical field which lead to demanding workloads for the integration and maintenance of data. Increased IT use for the patient-related requirements and generally a higher percentage of digitally available documents as opposed to the conventional paper-based ones also targeted. In order to achieve these goals the strategy included extension of the existing CIS, particularly the electronic health record (HER), into new functions including the economic, logistic and technical field, an improved knowledge management, introduction of a digital signature, integration of speech- and data communication, consolidation of the networks infrastructure and usage of mobile devices as much as investment into innovations such as the use of robotics and 3D visualization of images. (IT-Strategiebeirat der TILAK 2002, 3p). In a paper about the past, present and future of HIS, also regarding the TILAK example, Haux (2006, p. 278) concludes „that institution-centred hospital information systems are developing towards regional and global health information systems, with new, strongly extended functionalities and tasks." Further enforcing this statement is another 2006 paper proposing a three step strategy to move from paper-based discharge letters to a digital communication as is seen in the TILAK's EHR to foster a more efficient healthcare system with lesser redundancies (Schabetsberger et al. 2006).

The TILAK formulated a second strategy plan for the years 2008-2012 which still addressed similar goals to the first strategy such as extension of the KIS, full usage of the EHR, further developing the patient management and SAP functionality, more standardized processes, closer interfaces to public health networks including more options in eHealth and finally systems to analyse management and research. Current efforts are aiming to further digitalize the outpatient services and cooperating with Cerner to find solution for replacing the paper based fever charts (Stark; Schmied 6/1/2016, 38ff). With the beginning of this second strategy plan the TILAK spent about 2,8% of its over-all budget on IT, employing around 70 people in this sector (Lechleitner 2007).

In the state of Styria the KAGes is the biggest public healthcare provider and on its official webpages claims to be Styria's biggest employer with more than 17 000 employees in 28 locations (Steiermärkische Krankenanstaltengesellschaft m.b.H. 2016). Similar to the TILAK the KAGes developed a dedicated IT strategy for its hospitals with the EHR and interchangeability of patient data between different healthcare providers as central aspects. Arguably KAGes' early involvement in PACS and RIS related research under the supervision of key experts helped to move towards a more digitalized hospital (Leodolter, Kocever 2003). In spring 1998 the KAGes started a project in order to evaluate the most suitable vendor for a HIS that was supposed to be introduced into two pilot hospitals and subsequently distributed to all affiliated locations in 2003/04. An official tender yielded fifteen bids from the European market of which six offers seemed to fit the inclusion-criteria and consequently got chosen for further analysis. The KAGes went through substantial efforts by actually installing the software on test sites and having them evaluated by assessment teams consisting of different professions. Experiences with the products showed that they all were based on the same architectural concept, being highly customizable but none fitting the category "Standard software". Furthermore they all focused on medical documentation without clear answers how to solve connecting issues to administrative systems and revealing major drawbacks as in a lack of user support and guidance as much as not meeting the required standard for data protection (Gell et al. 2000, p. 152). Finally a strategic partner was found in a Vienna based IT company and their SAP based software IS-H/IS-H*MED which consequently resembled KAGes' CIS. The system named MEdical and nursing DOCumentation network of Styria or MEDOCS and since then became openMEDOCS, as standard interfaces with ELGA for better interoperability got developed. With more then 10 000 end-users openMEDOCS is now one of Europe's biggest HIS and was used in a study to find a holistic so-called *Total Workplace Usability* approach to enhance HIS applications front-end functionalities by considering the users' needs and requirements (Holzinger, Leitner 2005). According to the KAGes Business Report 2014 openMEDOCS has about 15 200 end-users with up to 4 800 people working in it simultaneously and an annual 1,79 million diagnoses being transferred into this CIS.

The KAGes claims to have an "Innovative Information management" as they opened a new 1000m² computation centre in 2014, have a highly functional IT helpdesk and 70% of patients using the possibility to check themselves into the hospitals with their insurance card on automated desks, reducing the administrative workload on front desks (Steiermärkische Krankenanstaltengesellschaft m.b.H. 2016, 25pp). Within the KAGes medical personnel can lock themselves into the CIS via a *Magic Card*, so that the user finds himself in the exact same application or file when changing work stations (Kocever; Schmied 6/6/2016, 28ff). Currently the KAGes enhances its own business warehouse to the more advanced SAP HANA technology that uses faster memory processing (Kocever; Schmied 6/6/2016, 51f)

6.2 Data Mining for Hospital Management

Within the relevant studies examining DM Applications on a HIS two major topics were represented heavily. With ten identified studies the most prominent exploration was into the prediction of Clinical Pathways, using event logs that by analysing the events and services a patient goes through enable a better insight into behavioural patterns related to a specific disease. Therefore such application may enhance a hospital's resource planning through an improved estimate of resources needed. Next six studies' topic was the analysis of length of stay (LOS), thus not using mere statistical methods to calculate the average days a patient with a particular diagnosis would remain hospitalized. Three studies looked into forecasting admission numbers specifically; two each were on information retrieval and infection control. It is important to note that the characteristics of a DM project may enable multiple applications. Once the according data is processed different models can be applied aiming to extract a variety of knowledge. For this reason the major focus of any DM application may intervene with other topics and therefore the data extracted, transformed and loaded could be utilized for multiple purposes. Other papers from the review were general studies on the use of DM in hospital settings or had a particular focus on e.g. risk management or resource planning.

6.2.1 Determination of Clinical Pathways

Clinical Pathways (CP) describe the processes a patient runs through based on a certain diagnosis, they include necessary routines and check-ups in form of reminders and recommendations on the treatment. Therefore they are to nurture evidenced-based medicine and have been used internationally since the 1980s (Kinsman et al. 2010). Optimizing the scheduled cares during a patient's hospitalization not only enhances the medical outcome and patient experience, it also supports an efficient clinical process management. CP habitually get developed manually by experienced medical staff (Tsumoto et al. 2015, p. 1204), thus a number of studies is on clinical process mining aiming to enhance or cultivate entire CP from collected data. Benefits from a deeper understanding of CP through mining techniques are also stated by the fact that the professionals and therefore "health-care organisations typically have an oversimplify and incorrect view of the actual situations in clinical pathways" (Huang et al.

2013, p. 112). Different approaches have been facilitated for the process mining of CP but most papers described the use of multiple clustering algorithms. Methods would also be influenced by the main focus of the process-mining. Orellana et al. (2015, p. 311) describe three kinds of process mining: *Discovery* to extract process models; *Conformance checking* to find deviations between models and the actual pathways taken and *Enhancement* or *Extension* to improve current models.

As a general technique in business analysis, process mining has been receiving an increasing amount of attention in the medical field, with the basic idea of using event logs to discover process models (Huang et al. 2013, p. 114). Event logs may be seen as a timestamp that relates information about a specific action processed in the HIS. It is advisable to include certain meta-data, e.g. process name or event log version, into the event log format as they are composed only once (García et al. 2015, p. 310). With specific tools these event logs may be extracted from the HIS and used for the process mining.

Rebuge & Ferreira (Rebuge, Ferreira 2012, p. 102) argue that "process mining for Business process analysis in health- care is a relatively unexplored field" and (Huang et al. 2013, p. 112) pointed out that "the use of traditional process mining techniques though successful in discovering clinical pathway models can prove inadequate in clinical pathway analysis". As successful and efficient applications Orellana et al. (2015, p. 311) highlight two examples from Emergency departments (ER) that could identify peak times for better resource allocation and detect eventualities such as missing information or unfinished orders within processes. With their own work the authors started towards the goal of creating process models that could be understood by non-experts and therefore support medical and organisational decision-making. They developed a component that ensures all traces processed in the HIS becoming stored in process event logs and claim that this is the first solution of this kind that is "designed to analyse processes in the systems HIS" (García et al. 2015, p. 313).

Rebuge & Ferreira (2012, 112ff) likewise were able to identify bottlenecks in the processes of an ER, that had its own EHR system. Interestingly the authors were able to "provide insight into the flow of healthcare processes, their performance, and their adherence to institutional guidelines" (Rebuge, Ferreira 2012, p. 115) in-house using the extensible software framework ProM. This platform is implemented in Java and downloadable free of charge, thus not causing additional costs for potential users. At least two of the studies mining CP facilitated *Petri Nets* within Pro M. Petri Nets or Place/Transition nets are a mathematical modelling language, graphically describing how activities are communicated and executed in complex software systems with multiple components across a network. Petri Nets like almost all modelling techniques have the drawback of becoming incredibly large with rising complexity of a system and then are difficult to validate on a lower level (Bobbio 2013, p. 138).

A study on temporal DM in HIS also was conducted by Tsumoto & Hirano (2011, p. 6) which were able to extract some analysis from Shimane University Hospital's stored data. For instance they found that 83% of all clinical orders were prescription, laboratory examination or reservation. They also derived a decision tree relating long-term follow up patients and their visit to clinical sections, showing that "number of orders of laboratory examinations is more than 40, 76.2% of the patients are visiting the department of internal" (Tsumoto, Hirano 2011, p. 5). Yet this particularly study was described as preliminary work and in 2014 Iwata, Hirano & Tsumoto proposed a *Similarity based Visualization* approach using clustering and multidimensional scaling to construct and maintain CP, which is seen as "a first step to [a] data-oriented approach into hospital management" (Iwata et al. 2014, p. 1114). Using a dual-, sample and attribute, clustering technique on nursing interventions for cataract operations Tsumoto et al. (2015, p. 1211) were finally able to not only extract the according CP but also discovered frequent nursing cares that were not associated before.

In terms of accuracy progresses with the evolution of more specialized mining algorithms certainly have been made. In 2003 Dart et al. constructed a model to predict a patient's path between the medical units based on association rules. K cross-validation showed that the model only classified 49,9% of the new patients accurately and the authors state that "medical data analysis using data-mining technologies has certainly to be developed" (Dart et al. 2003, 266f). About ten years later Eunhye et al. (2013, p. 45) derived a mega-process for outpatient cares from almost 700 000 event logs that had a 89.01% matching rate with the process model developed by clinical experts "regarding the total flow frequency". With such accuracy the authors claim that machine found patterns and processes could be used for improved resource allocation through the identification of bottlenecks and extraction of delays in processing (Eunhye et al. 2013, p. 48).

Finally Huang et al. (2013, p. 126) published a paper on the importance of clinical summaries and that they may also be derived through process-mining. The practical example was conducted on four kinds of cancer with the data from Zhejiang Huzhou Central hospital's HIS. The researchers suggested that using the same methodology also "useful business process summaries" could be constructed. In a following paper the same authors established the *Treatment Pattern Model* as a system generated from EHR that is capable of extracting critical behaviour patterns from CP. This model provides "predictive monitoring services in CP" including compliance checking, recommendations for treatment activities and is feasible for the pathway management (Huang et al. 2016, p. 227). Yet the model has a medical focus and the authors admitted certain drawbacks on the managerial side. The model could not anticipate or compensate for disturbances caused by inefficient operational proceedings. The allocation of insufficient or to many resources to a particular activity results in errors and variations such as postponed treatments and re-scheduled appointments (Huang et al. 2016, p. 237).

6.2.2 Length of Stay, Morbidity

Length of stay (LOS) is the amount of days a patient remains hospitalised. When the data is readily accessible a retrospective LOS is determined with simple statistical tools, yet "frequently used in health care as an indicator of resource utilization and cost" (Kraft et al. 2002, p. 6). In a broad perspective LOS "is the principle factor for the revenue of the hospital, whose distribution follows the log-normal distribution" (Tsumoto, Tsumoto 2006, p. 2). Huang et al. (2013, 111f) claim that commercial BI tools only provide an external view onto aggregated data in order to answer relatively simple questions such as average LOS and the costs related to a particular CP. Due to the analysis of event logs process mining on CP as seen in the preceding chapter would often enable a detailed look into LOS as well. It is notable that in a number of studies the original task was another one then extraction of LOS data but through the generic process such insights became possible. This fact is probably explained through the procedure of DM itself as in once the data got extracted, cleaned and formatted into an according (mining)format it is possible to run a variety of analysis. Young et al. (2003, p. 171) for example used a decision tree to identify the factors influencing patient's mortality most. The factors found to be most influencing were: LOS, the disease class, age group and the department from which the patient got discharged.

Of particular interest were those studies aiming to develop a model for precise prediction of LOS or its influencing factors. Such predictions became increasingly important as "awareness of factors and elements that determine LOS could promote the development of efficient clinical pathways and optimize resource utilization […], comprising the appropriate allocation of health care resources according to differences in patients' LOS along with considering patients' health status and social-demographic features "(Hachesu et al. 2013, p. 122).

One of the first studies regarding this topic is from 2002 when Kraft et al. (p. 5) used 525 patients with 1107 admissions to form 20 nursing diagnoses clusters that were then used to develop an ANN for the prediction of LOS. Out of an applied four different supervised learning models the *Radical Basis Function* was found to perform best with 77% predicted hospital days based on the patient's diagnosis and "one can remain confident that the model will provide reliable information when it is applied to real up-coming data to define the LOS of future patients" (Kraft et al. 2002, p. 7). Notably the data used was compromising eleven years from a veterans hospital specialized in spinal cord injury and to review the visualized data took the researches approximately 500 hours of time (Kraft et al. 2002, p. 5).

Young et al.'s (2003, p. 171) study on the decision tree model to understand influencing factors on patient morbidity highlights the benefit of understanding the characteristics of segmentation to develop strategies for continuous quality improvement, something that is not provided through logistic regressions. In their opinion further quality indicators should be analysed and DSS should "be well integrated with the hospital order communication system (OCS) to support concurrent review" (Young et al. 2003, p. 171).

In 2009 Belciug (p. 79) statistically evaluated an English database on stroke patients and used *agglomerative hierarchical clustering* for grouping. The particular algorithms were chosen as they aim to distinguish between different groups without being limited by the amount of expected clusters, with these clusters eventually corresponding to meaningful systematizations (Belciug 2009, p. 80). The possibility of other clustering algorithms performing better on this particular database is not ruled out so that further research is necessary (Belciug 2009, p. 84).

Compared to a decision tree and ANN, *support vector machines* (SVM) were found to be the most suitable classification algorithm to predict the mean LOS in patients with Coronary artery disease. SVM have "the highest forecasting accuracy among other DM algorithms. Today, this algorithm is becoming increasingly common in the medical and health field" (Hachesu et al. 2013, p. 128). The authors used 36 input variables to predict the LOS in cardiac patients. Statistical analysis showed a variety of influencing factors yielding that patients had a normal

LOS when they were free of preceding heart diseases, comorbidities and non-smokers. On the other hand patients using anticoagulant drugs could expect a prolonged LOS as much as married man in comparison to single men and women (Hachesu et al. 2013, p. 121).

Significant correlation between LOS and the variables surgery, gender, payment type, ward type, final diagnosis and age were found by Kalankesh et al. (2015, p. 213). The researchers used multiple regression analysis namely in Excel and the open-source software *RapidMiner* in order to analyse and visualize the data. Consequently it is possible to gain a deeper insight into data from a HIS if the mining process in terms of data extraction and cleansing is applied thoroughly and basic processing is done in readily available software. At least "mining" for basic reporting and to find significant correlations no sophisticated algorithms seem necessary. Yet the authors state that the development of research databases from HIS "appears to be inevitable if meaningful use of the data is to be achieved by researchers" and that "[m]ining trends in other administrative data and clinical data can add even more value to the health care management" (Kalankesh et al. 2015, p. 214).

6.2.3 Dedicated to Patient Forecasting and Business Intelligence Systems

Over the criteria used to identify the relevant studies it became obvious that they all may improve hospital management through better prediction and a consequently better resource allocation. While the investigations into CP and LOS seem well sophisticated methods to gain knowledge from DM HIS other studies do not particularly focus these topics, yet CP are often necessary in terms of Business Understanding. In an early work in 2000 Riano & Prado propose a system to help the medical and administrative communities "to model the hospital operation, to study the hospital costs and foresee the behaviour of new patients" (Riano, Prado 2000, 71f). Besides granting insight into the internal operations of a hospital, the methodology also revealed real costs based on the actual patient behaviour rather than statistical means gathered on a higher level.

Similar to the benefits of accurately predicting peak times and LOS the forecast of patient inflow may also support a hospital's strategic planning. Also an early work on using DM techniques on HIS in general is by Alapont et al. (2005). Their practical application focused on the forecast of average emergency admissions through different learning methods in the WEKA suite. The learning algorithms were performed upon *mineable view*s including seventeen attributes, internal factors, like number of admissions in previous days or the same month last year and external inputs like the occurrence of mayor sport events, holidays and meteorological data (Alapont et al. 2005, 5f). Compared to a mere statistical average the mean absolute error improved by 13.1 admissions per day for the linear regression model and 12.95 for the (decision) tree M5P model, allowing “hospital manager[s] to better adjust resources” (Alapont et al. 2005, p. 7).

To forecast outpatient visits to a hospital Hadavandi et al. (2012) used an AI model called *Clustering-Based Genetic Fuzzy System* (CGFS) offering a row of novel features that render it a “very powerful forecasting algorithm” (Hadavandi et al. 2012, 701ff). The authors’ evaluation of the model in a Taiwanese hospital yielded better results than other methods discussed in the literature, and its application to four Iranian hospitals was “very successful and promising” (Hadavandi et al. 2012, p. 710). The graphs in Figure 8 visualize CGFS forecasted values vs. the actual values for all cases in the Iranian hospitals, showing that the forecasts are capable of anticipating the tendencies in outpatient inflow.

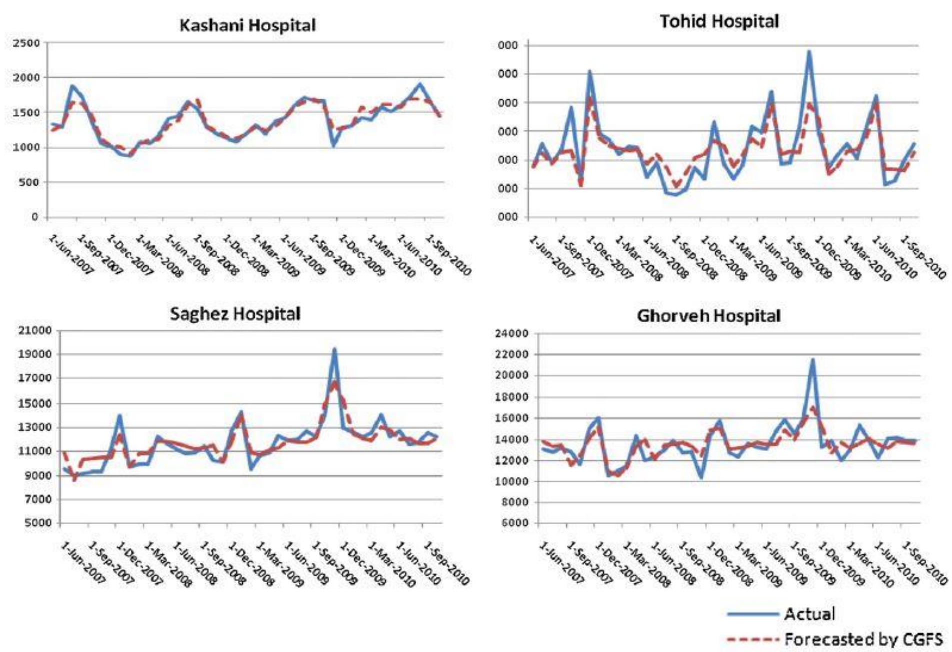

Figure 8: CGFS Forecasted Values vs. Actual Values for All Cases
(Source: E. Hadavandi et al. 2012, p. 709)

Oliveira et al. (2014) took a different approach by trying to forecast the number of discharges in order to be able to improve the bed management by consequently knowing the number of available beds for the subsequent weeks. The models were applied to four different wards in the Centre Hospital of Porto finding a SVM model to perform best. Results are shown in Figure 9 and the authors conclude that the suggested method makes it possible “to predict the number of weekly discharges of patients by using the number of patients discharged by day registered in the past“ (Oliveira et al. 2014, p. 1660).

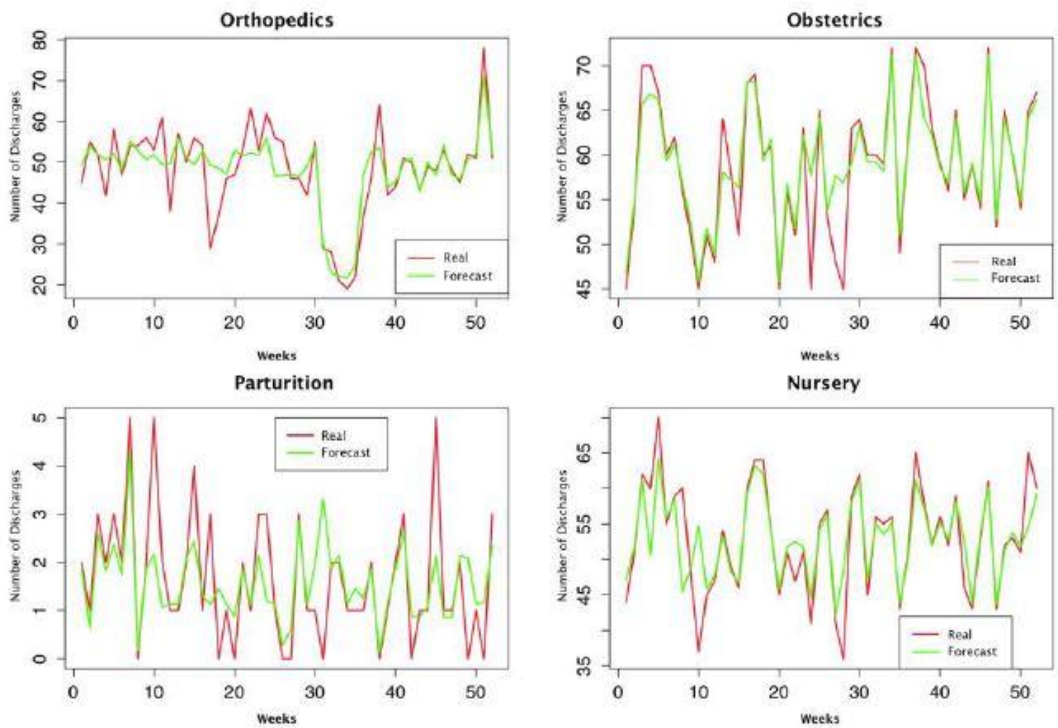

Figure 9: Number of Discharges Forecast vs. Actual Value (Source: Oliveira et al. 2014, p. 1659)

Managerial tools like dashboards and key metrics may not be utilized sufficiently in hospital environments. In Canada Grant et al. (2006) reported on a clinical DW that is updated from Sherbrooke University HIS every 24 hours. The authors constructed two prototype dashboards from the clinical data repository: The first was designed to analyse bed occupancy in the Emergency department, the second was aiming to ensure the evaluation of quality assurance within a biochemistry department. Now about ten years old the paper concluded with the authors anticipating a rapid expansion in the use and functionality of clinical data warehouses that could change the very core of hospital services as readily available compiled data would enable a continuous evaluation of the clinical practise (Grant et al. 2006).

The International Medical Centre of Japan implemented a universal three-layered IT system that manages Business Processes, Medical materials and medicine, Finances and the medical records consequently solving the “man, money, material and information [“] issues inherent in the costs of healthcare” (Akiyama 2007, p. 686). The so-called *Point of Action System* helped to improve the quality of care, reduced the error rate to nearly zero in certain cases and through higher efficiency in processes and lower logistical costs generated a “saving over four million dollars per year” (Akiyama 2007, p. 686). The underlying concept is that of each action within the hospital being captured and logged in operational track records. The information about all performed tasks, doctors and patients involved as much as the used material traceable via a barcode system gets compiled in the HIS and fed into a separate MIS each day at midnight. The MIS then processes the data delivering managerial information based on the previous day’s data at 6:00 am. Possible analysis are profit-and-loss calculations regarding the different medical divisions, patient groups and physicians as much as the calculation of costs by disease (Akiyama 2007, 688ff). Furthermore the tracking of individual drugs and medical supplies via a barcode system enables an effective inventory control. Replenishing of stocks can be automated and a reduction of the “inventory to a tenth of its previous level” was achieved, so that the hospital “was able to turn a profit after only one year” (Akiyama 2007, p. 689). The authors suggest facilitating time series analysis rather than comparing different departments given their divergent structure, e.g. the ER department being most unprofitable. The discussed Point of Action System resembles a good-practice example in terms of linking medical and logistical data to provide structured and accessible managerial information in terms of creating BI. A diversion between the HIS and MIS technically still exists but the proposed system seems to be an appropriate way of interfacing these two components improving the hospitals profitability and the treatment outcomes at the very same time.

In Taiwan Hao-Yun et al. (2016) used the *Design Science Research Methodology* in order to create a Hospital-based Business Intelligence System (HBIS) also including a three-tier server structure and an OLAP server with the conferring front-end tools. The used methodology was composed from six activities that may serve as a guideline for future development of HIS: problem definition, solution objectives, design and development as much as the demonstration, evaluation and communication of the findings (Hao-Yun et al. 2016, p. 496) and therefore the HBIS got developed based on the informational needs of the institution. The final tool enables flexible functions on a graphical user interface where analysts create need-specific queries by choosing “dimensions and attributes and drag them into the pivot report” (Hao-Yun et al. 2016, p. 499). Despite the challenges caused through administrational policies, the diversity of data sources and report objectives the HBIS was later evaluated to improve work efficiency. Furthermore the importance of providing a user friendly and well-structured end tool is highlighted, as “[t]he results also indicate that ease of use and interface usability are critical factors in improving work efficiency and streamlining the debugging process” (Hao-Yun et al. 2016, p. 500).

6.2.4 Other Studies

Lastly general studies on the purpose and techniques of DM in healthcare exist, sometimes combined with the presentation of a particular tool, e.g. the platform *DeepSee* in Jing-song et al. (2011). Also Rajalakshmi et al. (2011) see DSS as an emerging IT technology in healthcare, thus the introduction of many different technologies made it necessary to develop integrational projects such as HL7. All of the interview partners confirmed that such (medical) IT standards greatly improve the transition and possible analysis of data, yet the maintenance of those formats requires some work as well. HL7 and similar formats cannot be seen as Plug-and-Play Systems that effortlessly connect and operate.

Another field of applying DM techniques on HIS components is that of Infection Control. Already in 1998 Brossette et al. demonstrated an detection system based on association rules that efficiently identified interesting and previously unknown patterns in the occurrence of a particular bacterium (Brossette et al. 1998, p. 380). That interfacing a laboratory information system with an existing database for infection control can be a cost-saving alternative to otherwise expensive software systems was shown by Dao et al. (2008, p. 20). The proposed system provides graphical functions, automated calculation of infection rates and detects increases with the whole project been “designed and implemented at minimal cost for the hospital“ (Dao et al. 2008, p. 20). Computer based infection control including automated alerts are the topic of many specific studies and yet sophisticated solutions are available it remains unclear which systems get broadly facilitated in Austria.

Also the subject of risk mining gets discussed in the literature. The three processes of risk mining: risk detection, risk clarification and risk utilization are visualized in Figure 10:

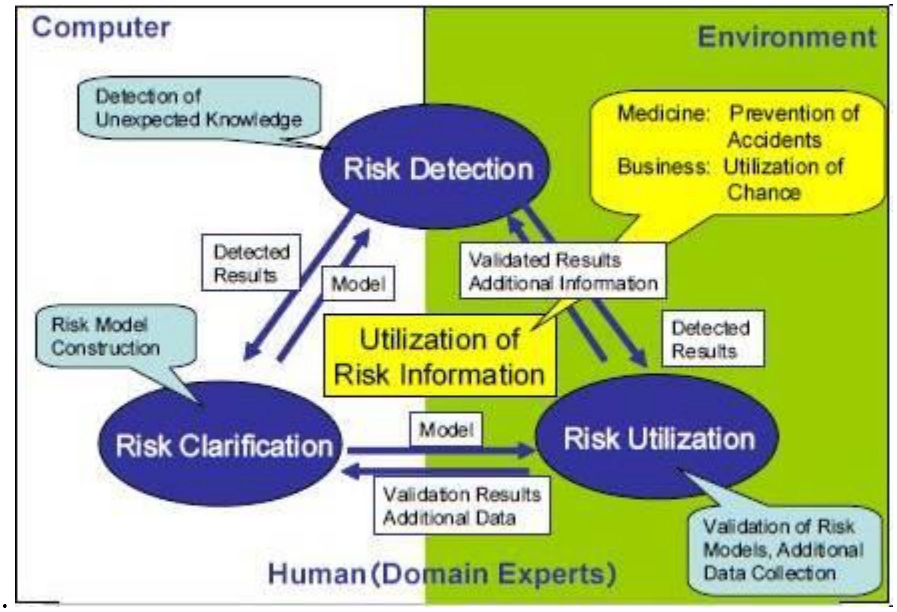

Figure 10 : Overview of the Risk Mining Process (Source: Tsumoto, Hirano, p. 3)

Mining risks is a problematic task as the methods rely on high frequencies while accidents and near-misses only happen at a very low rate and the according reports may hold some uncertainties due to the use of individual language (Tsumoto, Hirano, 2f). In a partial application three algorithms were used on 245 incident reports, including the type of near-miss, the patient factor, staff factor and the time of shift. The authors were able to establish rules such as: “If late-night and lack of checking, then medication errors occur” with a probability of 53,8% or in 8 out of 15 cases (Tsumoto, Hirano, p. 4). Based on the discovered rules a system of having two nurses prepare the medication after their official shift and a third responsible for another check was introduced with the “surprising” effect of a medication error reduction of 90% (Tsumoto, Hirano, p. 5).

Kurbjuhn & Schult (2010) suggest another approach to enhance a hospitals resource management based on data mining. They described a project to trace mobile medical equipment within a hospital via radio frequency identification (RFID) chips in order to improve planning and resource allocation. Their proposed method basically generates *event logs* for the equipment furnished with RFID allowing the analysis of the paths it was taken and therefore rate of usage. Exemplary the authors wrote that ambulatory infusion pumps resemble a bound value of 550 000 € in their study hospital and stocking levels could by optimized through better usage and the ease of locating the pumps (Kurbjuhn, B., Schult, R. 2010, p. 55). Yet the regarding papers are from 2006 and 2010 respectively and are meant to be discussion papers while the task of choosing valuable analysis goals and according algorithms would still have to be commenced in the project.

Finally the use of text-mining or semantical searches on the data stored in HIS offers a number of possibilities. The principle idea is to retrieve missed or missing information from the patients records and practical applications include the search for pressure ulcers as adverse events in hospitals (Gerdes, Hardahl 2013), extracting numerical values essential for the monitoring of Diabetes Mellitus and Arterial Hypertension Patients (Boytcheva et al. 2015) or mining discharge letters for lab tests, medications and diagnoses missed (Tchraktchiev et al. 2011). Some of the interview partners also saw semantical search as offering many future prospects, increasing a hospital's efficiency as much as the quality of treatment.

7. Limitations & Outlook

Limitations to this thesis are mostly due to the complexity and multitude of the systems forming a hospital’s IT structure. Within the institutions clinical and organisational systems have grown historically and a practical distinction between those two still is palpable. Austrian hospitals appear to be well set up with their clinical systems using standardized formats, digitalized imagery, a high proliferation of IT based nursing documentation and strong trends towards EHR and eHealth. Also Data warehousing and technical infrastructure is on a generally high level and at least within clinical networks dedicated business warehouses and software for analysis are commonly used. However a lack of knowledge still remains regarding the informational output of these individual solutions. The Business Intelligence gained may strongly vary and the level of pervasion of tools such as Managerial Dashboards or automated alerts persists to be unknown. It also remains uncertain how the software functions, if they do apply DM algorithms to identify unseen patterns or merely aggregate information. For this moment it is impossible to judge on the depth of analysis, if hospital managers are bound to base their decisions on retrospective reports or if their decision making gets supported by predictive explorations.

Regarding DM in healthcare the major fields of application could be identified. Nonetheless the amount of related studies seems immense and the conducted review by no means is comprehensive. Likewise to the processing of data it is hard to distinguish between a purely medical or business related use of the extracted information and interfacing all data sources appears to bear the highest potential for KDD. Consequently a big challenge is devising evocative and useful application models. Some of the identified examples generate conclusions and correlations that can be considered common knowledge for most practitioners and nurses even without confirmation by research. Hitherto the studies were of an explorative character and with improved interchangeability of data and wider spread utilization of DM the emergence of more expedient applications is very likely. For particular scientific issues best performing algorithms can be identified in the literature and their adequate introduction into the field depends on the hospital managers’ willingness to facilitate these methods. Possible benefits stem from improved predictions to support managerial decision making and strategic planning and due to Austria’s high standards in HIS could relatively easy be obtained.

Further research on this topic should aim to identify the best approaches to facilitate data stored in a HIS for BI purposes. It would be sensible to develop a request catalogue for the upper hospital management according to which current solutions could be assessed and benchmarked. Being able to evaluate and scale different hospital BI solutions could help to define the procedures and structure for an efficient set-up or improve current systems. Preliminary studies are available and suitable instruments may be developed by adapting the HIS Monitor to a managerial level or the use of the Nolan maturity index.

8. Publication Bibliography

Agfa-Gevaert Group (2016): AGFA HealthCare - HYDMedia. Online. Available online at http://www.agfahealthcare.com/germany/de/main/products_services/dms/index.jsp, checked on 6/26/2016.

Akiyama, M. (2007): Risk Management and Measuring Productivity with POAS – Point of Act System. A Medical Information System as ERP (Enterprise Resource Planning) for Hospital Management. In *Methods of Information in Medicine* 46 (6), pp. 686–693, checked on 5/4/2016.

Alapont, J.; Bella-Sanjuán, A.; Ferri, C.; Hernánd-Orallo, J.; Llopis-Llopis, J. D.; Ramírez-Quintana, M. J. (2005): Specialised Tools for Automating Data Mining for Hospital Management. Valencia, Spain. Available online at http://users.dsic.upv.es/~abella /papers/HIS_DM.pdf, checked on 1/11/2016.

Ammenwerth, E.; Ehlers, F.; Hirsch, B.; Gratl, G.(2007): HIS-Monitor: an approach to assess the quality of information processing in hospitals. In *International journal of medical informatics* 76 (2-3), pp. 216–225.

Ammenwerth, E.; Rauchegger, F.; Ehlers, F.; Hirsch, B.; Schaubmayr, C. (2011): Effect of a nursing information system on the quality of information processing in nursing: An evaluation study using the HIS-monitor instrument. In *International journal of medical informatics* 80 (1), pp. 25–38.

Avin, M. (2008): Asset Management Data Warehousing Data Modelling. Doctor. School of Engeneering Systems. Available online at http://eprints.qut.edu.au/19310/1/ Avin_Mathew_Thesis.pdf checked on 3/30/2016.

Baca-García E.; Perez-Rodriguez, M. M.; Basurte-Villamor, I.; Saiz-Ruiz, J.; Leiva-Murillo, J. M.; Prado-Cumplido, M. de et al. (2006): Using Data Mining to Explore Complex Clinical Decisions: A Study of Hospitalization After a Suicide Attempt. In *Journal of Clinical Psychiatry* 67 (7), pp. 1124–1132, checked on 2/12/2016.

Belciug, S. (2009): Patients length of stay grouping using the hierarchical clustering algorithm. In : Annals of University of Craiova, Math. Comp. Sci. Ser., 36(2), pp. 79–84. Available online at http://inf.ucv.ro/~ami/index.php/ami/ article/viewFile/288/279, checked on 2/26/2016.

Biron, P.; Metzger, M. H.; Pezet, C.; Sebban, C.; Barthuet, E.; Durand, T. (2014): An information retrieval system for computerized patient records in the context of a daily hospital practice: the example of the Léon Bérard Cancer Center (France). In *Applied clinical informatics* 5 (1), pp. 191–205.

Bobbio, A. (2013): System Modelling with Petri Nets. In A. G. Colombo, A. Saiz de Bustamante (Eds.): Systems Reliability Assessment. Proceedings of the Ispra Course, Madrid, Spain, September 19-23, 1988, vol. 6. Dordrecht: Springer Verlag (ISPRA Courses), pp. 103–143. Available online at https://www.fd.cvut.cz/department/k611/PEDAGOG/THO_A/A_soubory/ miopetrinet.pdf, checked on 5/31/2016.

Boytcheva, S.; Angelova, G.; Angelov, Z.; Tcharaktchiev, D. (2015): Text Mining and Big Data Analytics for Retrospective Analysis of Clinical Texts from Outpatient Care. In *Cybernetics and Information Technologies* 15 (4).

Breitschwerdt, E. (2011): Wissensmanagement, Business Intelligence und Business Analytics für das Wissenskrankenhaus. In *Krankenhaus - IT Journal* 6/2011, checked on 1/9/2016.

Brossette, S. E.; Sprague, A. P.; Hardin, J. M.; Waites, K. B.; Jones, W. T.; Moser, S. A. (1998): Association Rules and Data Mining in Hospital Infection Control and Public Health Surveillance. In *Journal of the American Medical Informatics Association* 5 (4), pp. 373–381.

Cano Giner, J. L.; Fernandez, V.; Díaz Boladeras, M. (2009): Framework for the analysis of executive information systems based on the perceived usefulness and the perceived ease of use. In *Intangible Capital* 5 (4).

Chapman, P.; Clinton, J.; Kerber, R.; Khabaza, T.; Reinartz, T.; Shearer, C.; Wirth, R. (2000): CRISP-DM 1.0. Step-by-step data mining guide. The CRISP-DM consortium. Available online at https://the-modeling-agency.com/crisp-dm.pdf, checked on 4/11/2016.

Chen, H.; Chian, R. H. L.; Storey, V. C. (2012): Business Intelligence and Analytics. From Big Data to Big Impact. In *MIS Quarterly* Vol. 36 No. 4 (Special Issue: Business Intelligence Research), pp. 1165–1188, checked on 12/19/2015.

Chmielarz, W.; Tasic, V.; Szumski, O. (n.d.): Management Information Systems. University of Warsaw. Faculty of Management, n.d. Available online at www.wz.uw.edu.pl/pracownicyFiles/id4236-mis4ibp2014.ppt, checked on 4/2/2016.

Dao, T. K.; Zabaneh, F.; Holmes, J.; Disrude, L.; Price, M.; Gentry, L. (2008): A practical data mining method to link hospital microbiology and an infection control database. In *American Journal of Infection Control* 36 (3), pp. S18-S20.

Dart, T.; Cui, Y.; Chatellier, G.; Degoulet, P. (2003): Analysis of hospitalised patient flows using data-mining. In *Studies in Health Technology and Informatics* 95, pp. 263–268, checked on 5/4/2016.

DeGruy, K. B. (2000): Healthcare Applications of Knowledge Discovery in Databases. In *Journal of Healthcare Information Management* vol. 14 (no. 2, Summer 2000), checked on 12/20/2015.

Durairaj, M.; Ranjani, V. (2013): Data Mining Applications In Healthcare Sector. A Study. In *International Journal of Scientific & Technology Research* Volume 2, (10, October 2013), checked on 1/11/2016.

Eltabakh, M. (2012): OLAP and Data Mining. Worcester Polytechnic Institute, 2012. Available online at http://web.cs.wpi.edu/~cs561/s12/Lectures /IntegrationOLAP/ OLAPandMining.pdf, checked on 4/1/2016.

Eunhye, K.; Seok, K.; Minseok, S.; Seongjoo, K.; Donghyun, Y.; Hee, H.; Sooyoung, Y. (2013): Discovery of outpatient care process of a tertiary university hospital using process mining. In *Healthcare informatics research* 19 (1), pp. 42–49.

Fayyad, U. M., Piatetsky-Shapiro, G.; Smyth, S. (1996): From Data Mining toKnowledge Discovery in Databases. In *AI Magazine* Fall. Available online at http://www.csd.uwo.ca/faculty/ling/cs435/fayyad.pdf, checked on 4/10/2016.

García, A. O.; Alfonso, D. P.; Larrea Armenteros, O. U. (2015): Analysis of Hospital Processes with Process Mining Techniques. In *Studies in Health Technology and Informatics* volume 216, Part 1-2, checked on 5/4/2016.

Gell, G.; Madjaric, M.; Leodolter, W.; Köle, W.; Leitner, H. (2000): HIS purchase projects in public hospitals of Styria, Austria. In *International journal of medical informatics* 58–59, pp. 147–155.

Gerdes, L. U.; Hardahl, C. (2013): Text Mining Electronic Health Records to Identify Hospital Adverse Events. In *Studies in Health Technology and Informatics* 192 (MEDINFO).

Gluchowski, P.; Gabriel, R.; Dittmar, C. (2001): Business Intelligence. Konzepte, Technologien, Einsatzbereiche. In *HMD - Praxis der Wirtschaftsinformatik* 38 (222), pp. 5–15, checked on 12/3/2015.

Gluchowski, P.; Gabriel, R.; Dittmar, C. (2008): Management Support Systeme und Business Intelligence. Berlin Heidelberg: Springer-Verlag, checked on 12/3/2015.

Govaert, J. A.; van Bommel, A. C. M.; van Dijk, W. A.; van Leersum, N. J.; Tollenaar, R. A.; Wouters, M. W. (2015): Reducing healthcare costs facilitated by surgical auditing: a systematic review. In *World journal of surgery* 39 (7), pp. 1672–1680.

Grant, A.; Moshyk, A.; Diab, H.; Caron, P.; De Lorenzi, F.; Bisson, G. (2006): Integrating feedback from a clinical data warehouse into practice organisation. In *International Journal of Medical Informatics* 75 (3), pp. 232–239.

Hachesu, P. R.; Ahmadi, M.; Alizadeh, S.; Sadoughi, F. (2013): Use of data mining techniques to determine and predict length of stay of cardiac patients. In *Healthcare informatics research* 19 (2), pp. 121–129.

Hadavandi, E.; Shavandi, H.; Ghanbari, A.; Abbasian-Naghnehd, S. (2012): Developing a hybrid artificial intelligence model for outpatient visits forecasting in hospitals. In *Applied Soft Computing* 12, pp. 700–711.

Hao-Yun, K.; Min-Chun, Y.; Mehedi, M.; Wen-Hsiung, W.; Li-Ju, Y.; Yen-Chun, W. J. (2016): Design and evaluation of hospital-based business intelligence system (HBIS). A foundation for design science research methodology. In *Computers in Human Behavior* 62, pp. 495–505.

Holzinger, A.; Leitner, H. (2005): Lessons from Real-Life Usability Engineering in Hospital: From Software Usability to Total Workplace Usability. In : Empowering Software Quality: How can Usability Engineering reach these goals? 1st Usability Symposium, Vienna, Austria, 8 November 2005, pp. 153–160.

Huang, Z.; Dong, W.; Ji, L.; Duan, H. (2016): Predictive monitoring of clinical pathways. In *Expert Systems with Applications* 56, pp. 227–241.

Huang, Z.; Lu, X.; Duan, H.; Fan, W. (2013): Summarizing clinical pathways from event logs. In *Journal of Biomedical Informatics* 46 (1), pp. 111–127.

Hübner, U.; Liebe, J. D.; Hüsers, J.; Thye, J.; Egbert, N.; Hackl, W.; Ammenwerth, E. (2015): IT-Report Gesundheitswesen. Schwerpunkt Pflege im Informationszeitalter. In *Schriftenreihe der Hochschule Osnabrück*, pp. 1–150, checked on 6/7/2016.

Hübner, U.; Ammenwerth, E.; Flemming, D.; Schaubmayr, C.; Sellemann, B. (2010): IT adoption of clinical information systems in Austrian and German hospitals: results of a comparative survey with a focus on nursing. In *BMC medical informatics and decision making* 10, p. 8.

Hummeltenberg, W. (2014): Business Intelligence. Edited by N. Gronau, J. Becker, K. Kurbel, E. Sinz, L. Suhl. Online (Enzyklopaedie der Wirtschaftsinformatik - Online Lexikon). Available online at http://www.enzyklopaedie-der-wirtschaftsinformatik.de /lexikon/daten-wissen/Business-Intelligence, checked on 4/1/2016.

Ikart, E. M. (2005): Executive Information Systems and the Top-Officer's Roles: An Exploratory Study of User-Behaviour Model and Lesson's Learnt. In *Australian Journal of Information Systems* 13 (1, September). Available online at http://ro.uow.edu.au /cgi/viewcontent.cgi?article=2917&context=commpapers, checked on 4/6/2016.

IT-Strategiebeirat der TILAK (Ed.) (2002): TILAK IT-Strategie. Informationstechnologie im Dienste von Patientenversorgung und medizinischer Forschung. With assistance of Ass.-Prof. Dr. E. Ammenwerth (Redaktion), Univ.-Prof. Dr. R. Haux, Univ.-Lekt. Dr. G. Lechleitner, Univ.-Prof. Dr. K.P. Pfeiffer, Dir. C. Triendl, Dr. R. Vogl., checked on 5/4/2016.

Iwata, H.; Hirano, S.; Tsumoto, S. (2014): Construction of Clinical Pathway based on Similarity-based Mining in Hospital Information System. In *Procedia Computer Science* 31, pp. 1107–1115.

Jahanbakhsh, M.; Sharifi, M.; Ayat, M. (2014): The status of hospital information systems in Iranian hospitals. In *Acta informatica medica : AIM : journal of the Society for Medical Informatics of Bosnia & Herzegovina : časopis Društva za medicinsku informatiku BiH* 22 (4), pp. 268–275.

Jing-song, L.; Hai-yan, Y.; Xiao-guang, Z. (2011): Data Mining in Hospital Information System. NewFundamental Technologies in Data Mining. Edited by Prof. Kimito Funatsu (Ed.). InTech (978-953-307-547-1,), checked on 1/11/2016.

Kalankesh, L. R.; Pourasghar, F.; Jafarabadi, M. A.; Khanehdan, N. (2015): Depiction of Trends in Administrative Healthcare Data from Hospital Information System. In *Materia socio-medica* 27 (3), pp. 211–214.

Kinsman, L.; Rotter, T.; James, E.; Snow, P.; Willis, J. (2010): What is a clinical pathway? Development of a definition to inform the debate. In *Bio Med Central medicine* 8, p. 31.

Koh, H. C.; Tan, G. (2005): Data Mining Applications in Healthcare. In *Journal of Healthcare Information Management* Vol.19, (No.2, Spring 2005), checked on 2/1/2016.

Kooperation für Transparenz und Qualität im Gesundheitswesen (2015): KTQ - Katalog Version 5.0. Kohlhammer Verlag (ISBN: 9783170300712). Available online at http://www.lwl.org/klinik_muenster_download/pdf/KTQ_5_0_Katalog_Kap_9.pdf, checked on 5/18/2016.

Kraft, M. R.; Desouza, K. C.; Androwich, I. (2002): Data Mining in Healthcare Information Systems: Case Study of a Veterans'Administration Spinal Cord Injury Population. Edited by IEEE Computer Society. Proceedings of the 36th Hawaii International Conference on System Sciences, checked on 2/12/2016.

KTQ International (2016): Official Webpage. Online. Available online at http://www.ktq-international.com/index.php?id=12, checked on 5/18/2016.

Kundu, A.; Pallavi (2015): Medical Data Mining: A Review. In *International Journal of Advanced Technology in Engineering and Science* Volume No 03, March (Special Issue No. 01), Available online at http://ijates.com/images /short_pdf/1426662128_699.pdf. checked on 1/11/2016

Kurbjuhn, B., Schult, R. (2010): Potenzial des Data Mining für Ressourcenoptimierung mobiler Geräte im Krankenhaus. LWA, checked on 2/11/2016.

Lechleitner, G. (2007): IT-Strategie der Tilak 2007-2012. Tilak. Online (12. Fachtagung „Praxis der Informationsverarbeitung in Krankenhaus und Versorgungsnetzen“). Available online at http://services.informatik.hs-mannheim.de/~kis/ lu/vortraege/Lechleitner_21062007.pdf, checked on 5/16/2016.

Leodolter, W.; Kocever, K. (2003): PACS as a driver for integrating healthcare systems. In *International Congress Series* 1256, pp. 910–914.

Lloyd, J. (2011): Identifying Key Components of Business Intelligence Systems and Their Role in Managerial Decision making. Master's Dissertation. University of Oregon. Interdisciplinary Studies Program, checked on 3/29/2016.

Menken, I. (2013): Data Warehousing Complete Certification Kit - Core Series for IT: Emereo Publishing. Available online at https://books.google.cz/books?id=TF-9CgAAQBAJ. checked on 12/2/2016.

Miller, K. (2012): Big Data Analytics In Biomedical Research. In *Biomedical Computation Review* (Winter 2011/2012), checked on 2/1/2016.

Mwachofi, A.; Al-Assaf, A. F. (2011): Health Care Market Deviations from the Ideal Market. In *Sultan Qaboos Univ Med J.* 11 (3), pp. 328–337. Available online at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3210041/, checked on 2/28/2016.

Mykityshyn, M. G. (2007): Assessing the Maturity of Information Architectures for Complex Dynamic Enterprise Systems. Dissertation (PhD). Institue of Technology, Georgia. Available online at https://books.google.cz/books?id=wHE2t41Ot6cC&pg=PA111&lpg=PA111&dq =Nolan+Maturity+Matrix&source=bl&ots=mgYA5FH5Cb&sig =SCUxP_SxEwxuyeb10W2fZk0CglQ&hl=en&sa=X&ved=0ahUKEwiN6qrJusrMAhWDVRQ KHSHpD1MQ6AEINTAE#v=onepage&q=Nolan%20Maturity%20Matrix&f=false, checked on 5/8/2016.

OLAP Council (1997): White Paper. Available online at http://www.symcorp.com /downloads/OLAP_CouncilWhitePaper.pdf, checked on 4/1/2016.

Oliveira, S.; Portela, F.; Santos, M. F.; Machado, J.; Abelha, A. (2014): Hospital bed management support using regression data mining models. Edited by International Work-Conference on Bioinformatics and Biomedical Engeneering. Granada. Available online at http://iwbbio.ugr.es/ 2014/papers/ IWBBIO_2014_paper_184.pdf, checked on 2/12/2016.

Olson, D. L.; Delen, D. (2008): Advanced Data Mining Techniques: Springer Berlin Heidelberg. Available online at https://books.google.cz/books?id=2vb-LZEn8uUC, checked on 4/11/2016.

Olszak, C. M.; Ziemba, E. (2006): Business Intelligence Systemsin the Holistic Infrastructure DevelopmentSupporting Decision-Making in Organisations. In *Interdisciplinary Journal of Information, Knowledge, and Management* 1, pp. 47–58. Available online at http://www.ijikm.org/Volume1/IJIKMv1p047-058Olszak19.pdf, checked on 4/27/2016.

Olszak, C. M.; Batko, K. (2012): The Use of Business Intelligence Systems in Healthcare Organizations in Poland. In. Federated Conference on Computer Science and Information Systems: IEEE, pp. 969–976, checked on 11/4/2015.

O'Neill, D. (2011): The potential role of Maturity Models in an assessment of Hospital I.T. Capability in Ireland. Dissertation. University of Dublin. Available online at https://cs.tcd.ie/postgraduate/healthInformatics/assets/pdfs/CMM_HITCAP%20assessment_Final_DERMOTONEILL.pdf, checked on 6/8/2016.

Österreichische e-Health Initiative (EHI) (Jänner 2007): Empfehlung für eine österreichische e-Health Strategie. Eine Informations- und Kommunikationsstrategie für ein modernes österreichisches Gesundheitswesen. Edited by K. P. Pfeiffer. Österreichische e-Health Initiative. Available online at https://www.i-med.ac.at/msig/service/oeehealth_strategie.pdf, checked on 5/14/2016.

Parvez, A.; Saqib, Q.; Afser Rizvi, S. Q. (2015): Techniques of Data Mining In Healthcare. A Review. In *International Journal of Computer Applications* 120 (15), pp. 38–50.

Paré, G.; Jaana, M.; Sicotte, C. (2009): Exploring Health Information Technology Innovativeness and its Antecedents in Canadian Hospitals. In *Methods of Information in Medicine* 49 (1), pp. 28–36.

Potocnik, M. (2006): 10 Jahre Leistungsorientierte Krankenanstaltenfinanzierung in Österreich. Master. Medizinische Universität, Graz, checked on 2/28/2016.

Prather, J. C.; Lobach, D. F.; Goodwin, L. K.; Hales, J. W.; Hage, M. L.; Hammond, W. E. (1997): Medical data mining: knowledge discovery in a clinical data warehouse. In *Proc AMIA Annual Fall Symposium*. Available online at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2233405/pdf/procamiaafs00001-0140.pdf, checked on 12/9/2015.

Qualitas (2009): Berater im Gesundheitswesen auf einen Blick. With assistance of Schaffler-Verlag. Qualitas Management Consulting CEE. Available online at http://www.schaffler-verlag.com/gesundheitswirtschaft/docs/Berater_ Qualitas.pdf, checked on 2/11/2016.

Rajalakshmi, K.; Chandra Mohan, S.; Babu, S. D. (2011): Decision Support System in Healthcare Industry. In *International Journal of Computer Applications* 26 (9), pp. 42–44.

Raphael, H. (2014): Business Intelligence im Krankehausmanagement. Herausforderungen an Kliniken im DRG Zeitalter. Wiesbaden: Springer Fachmedien.

Rebuge, Á.; Ferreira, D. R. (2012): Business process analysis in healthcare environments. A methodology based on process mining. In *Information Systems* 37 (2), pp. 99–116.

Reinhold H. (2006): Hospital information systems--past, present, future. In *International journal of medical informatics* 75 (3-4), pp. 268–281.

Riano, D.; Prado, S. (2000): A data-mining alternative to model hospital operations: clinical costs and predictions. In *Data Mining II*, pp. 63–72, checked on 2/29/2016.

Sakshi, B. (2012): Things to know about ETL in Business Intelligence. Paxcel Labs. n.p. Available online at http://paxcel.net/blog/things-to-know-about-etl-in-business-intelligencebi/, checked on 3/29/2016.

Schabetsberger, T.; Ammenwerth, E.; Andreatta, S.; Gratl, G.; Haux, R.; Lechleitner, G.l. (2006): From a paper-based transmission of discharge summaries to electronic communication in health care regions. In *International journal of medical informatics* 75 (3-4), pp. 209–215.

Schmied, M. (6/1/2016): Transcript #2. Interview with C. Stark. Prague - Innsbruck. Partial transcript; Audio Recordings.

Schmied, M. (6/3/2016): Transcript #1. Interview with E. Ammenwerth. Prague - Hall in Tirol. Partial transcript; Audio Recordings.

Schmied, M. (6/6/2016): Transcript #3. Interview with K. Kocever. Prague - Graz. Partial transcript; Audio Recordings.

Schult, R.; Kurbjuhn, B. (2006): Potenzial des Data Mining fur Ressourcenoptimierung mobiler Geräte im Krankenhaus. With assistance of Magdeburg Otto-von-Guericke-Universitat, checked on 2/4/2016.

Serova, E. (2012): Enterprise Information Systems of new Generation. 15 Issue 1 2012, (pp116 - 126), available online at www.ejise.com. In *The Electronical Journal Information Systems Evaluation* 15 (1), checked on 4/10/2016.

Steiermärkische Krankenanstaltengesellschaft m.b.H. (2016): KAGes - Das Gesundheitsunternehmen der Steiermark. Online. Available online at http://www1.kages.at/jobs-bildung, checked on 5/17/2016.

Tchraktchiev, D.; Angelova, G.; Boytcheva, S.; Angelov, Z.; Zacharieva, S. (2011): Completion of structured patient descriptions by semantic mining. In *Studies in Health Technology and Informatics* 166, pp. 260–269.

Thomsen, C.; Pedersen, T. B. (2008): A Survey of Open Source Tools for Business Intelligence. Institut for Datalogi, Aalborg Universitet. Aalborg (Technical Report; No. 23), checked on 2/3/2016.

Tirol Kliniken (n.d.): Leistungsbericht der Tirol Kliniken GmbH 2014. With assistance of Abteilung Finanzen und Beteiligungscontrolling. Tilak. Online. Available online at https://www.tirol-kliniken.at/page.cfm?vpath=index, checked on 5/16/2016.

Toba, S.; Tomasini, M.; Yang, Y. H.: Supply Chain Management in Hospital: A CaseStudy. In *California Journal of Operations Management* 6 (1), pp. 49–55. Available online at http://www.csupom.org/publications/2008/08-7.pdf, checked on 2/28/2016.

Tolan, N. V.; Parnas, M. L.; Baudhuin, L. M.; Cervinski, M. A.; Chan, A. S.; Holmes, D. T. et al. (2015): "Big Data" in Laboratory Medicine. In *Clinical Chemistry* 61 (12), pp. 1433–1440.

Tomar, D.; Agarwal, S. (2013): A survey on Data Mining approaches for Healthcare. In *IJBSBT* 5 (5), pp. 241–266.

Tsumoto, S.; Hirano, S. (Eds.) (n.d.): Temporal Data Mining in Hospital Information Systems. 11th International Conference on Cognitive Informatics & cogitive Computing. Kyoto, 22-24 Aug. 2012. IEEE Computer Society. Available online at http://ultimavi.arc.net.my/ave/IJCAI2011/Final/ijcai2011_h2.pdf, checked on 4/9/2016.

Tsumoto, S.; Hirano, S.: Data Mining for Risk Management in Hospital Information Systems. In : National Science Foundation Symposium on Next Generation, vol. 2007. Available online at http://www.csee.umbc.edu/ ~hillol/NGDM07/abstracts/poster/ STsumoto.pdf, checked on 2/12/2016.

Tsumoto, S.; Hirano, S. (2011): Temporal Data Mining in Hospital Information Systems. In *6th International Workshop on Chance Discovery (IWCD6)*. Available online at http://ultimavi.arc.net.my/ave/IJCAI2011/Final/ijcai2011_h2.pdf, checked on 2/12/2016.

Tsumoto, S.; Hirano, S.; Iwata, H. (2015): Mining Schedule of Nursing Care Based on Dual-Clustering. In *Procedia Computer Science* 55, pp. 1203–1212.

Tsumoto, Y.; Tsumoto, S. (2006): Construction of Statistical Models for Hospital Management. The R User Conference. Wien. Available online at https://www.r-project.org/conferences/useR-2006/Abstracts/Tsumoto+Tsumoto.pdf, checked on 2/12/2016.

Velez Lapão, L. (2007): Survey on the Status of the Hospital Information Systems in Portugal. In *Methods Inf Me*d 46(4):493-9

Whymark, G. K.: Management Support Systems. Principles and Concepts. In *Health Informatics in Australia*. Available online at http://www.achi.org.au /docs/HNI_Book/Chapter_11.pdf, checked on 6/24/2016.

Whymark, G. K. (1991): Development of the EIS Concept and its Implementation in the RAN. In *Australian Computer Journal* 23 (3).

Williams, S.; Williams, N. (2003): The Business Value of Business Intelligece. In *Business Intelligence Journal* (Fall). Available online at http://decisionpath.com/wp-content/uploads/2010/12/The-Business-Value-of-BI.pdf, checked on 2/26/2016.

Xie, Y.; Schreier, G.; Chang, D. C. W.; Neubauer, S.; Liu, Y.; Redmond, S. J.; Lovell, N. H. (2015): Predicting Days in Hospital Using Health Insurance Claims. In *IEEE Journal of Biomedical and Health Informatics* 19 (4), pp. 1224–1233.

Young, M. C.; Hye, S. K.; Kwan, C. T.; Hyun, J. P.; Seung, H. H. (2003): Analysis of healthcare quality indicator using data mining and decision support system. In *Expert Systems with Applications* 24 (2), pp. 167–172.