University of Tartu

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Institute of Computer Science
Software Engineering Curriculum

Agustín Guerrero

Extending the *L** Process Mining Model with Quality Management and Business Improvement Tools and Techniques

Masters' Thesis (30 ECTS)

Supervisor: Marlon Dumas, PhD

August, 2015

Extending the L Process Mining Model with Quality Management and Business Improvement Tools and Techniques*

Abstract

The purpose of this thesis is to determine whether is possible to expand the *L* life-cycle model* with Six Sigma's *DMAIC* model, the ISO 9001:2008 Quality Management System, and business improvement frameworks like the Baldrige Criteria for Performance Excellence for Business and Nonprofit™, and the European Foundation for Quality Management Excellence Model™. The work related to the Process Mining project where the *L* life-cycle model* was expanded with Six Sigma's *DMAIC* model has been conducted in an Italian IT Company with data from company's Help Desk and Software Quality Assurance operations. The work conducted in the company pursues in proving that the *DMAIC* cycle can provide an expanded framework for the *L* life-cycle model* in all of its stages while employing state of the art Process Mining techniques and Process Mining software.

Keywords: Process mining, Business process management, Six Sigma, DMAIC, ISO 9001, L* model, EFQM, Baldrige Program, KPI

L protsessikavandamismudeli laiendamine kvaliteedijuhtimise ning ariparandustööriistade ja -tehnikateg*

Resümee

Selle lõputöö ülesandeks on leida, kas *L* elutsükli mudelit* on võimalik laiendada Six Sigma *DMAIC* mudeli, ISO 9001:2008 kvaliteedijuhtimissüsteemi ja äriparandusraamistikega nagu Baldrige Criteria for Performance Excellence^{TM} äri ja mittetulundusühingutele ning European Foundation for Quality Management Excellence Model^{TM}. Protsessikaevandamisprojektiga, mille *L* elutsükli mudel* laiendati Six Sigma *DMAIC* metoodikaga, seotud töö viidi läbi Itaalia IT firmas kasutades andmeid firma abilauast ning tarkvara kvaliteedikontrolli tegevustest. Firmas läbi viidud töö näitab, et *DMAIC* tsükkel saab pakkuda laiendatud raamistikku *L* elutsükli mudelile* selle kõikides staadiumites kasutades tänapäevaseid protsessikaevandamistehnikaid ning tarkvara.

Märksõnad: Protsesside kaeve, äri protsesside juhtimine, Six Sigma, DMAIC, ISO 9001, L* mudel, EFQM, Baldrige Program, KPI

Acknowledgements

I would like to express my gratitude to Prof. Marlon Dumas for his dedication and guidance during the course of two years while enrolled in the Masters at University of Tartu. I also would like to thank Fabrizio Maggi, Abel Armas, and Luciano Garcia for their dedication and support during my studies.

I also would like to express my sincere appreciation to my father, Agustin Guerrero Sr., for his unconditional support during my studies in Estonia.

My gratitude also goes to Prof. Alessandro Sperduti, Daniele Turato, and Alessandro Berti from University of Padova for hosting me during my internship.

An acknowledgement is also extended to the staff at the Faculty of Mathematics and Computer Science as well as the team at the International Student Service for their support during my studies in Tartu and abroad.

Finally, I would like to express my gratitude for a free tuition Masters program granted by the Republic of Estonia, the DoRa and IT Academy scholarships, and the Erasmus+ traineeship abroad grant.

A.G. August, 2015

Republic of Estonia
Government

European Union
European Social Fund

Investing in your future

Co-funded by the
Erasmus+ Programme
of the European Union

Study IT in .ee
sponsored by Skype™

DoRa

For education and research in Estonia
ARCHIMEDES

Contents

List of Figures and Tables

CHAPTER 1

Introduction

Process Mining is a relatively new discipline which is the union of data mining and business process modeling[1], W. M. P. van der Aalst and et al. state that Process Mining can be seen as a technology which goal is to extract explicit process models from event logs[2]. Anne Rozinat states that Process Mining is the union between Data Mining and Process Modeling, that is, Process Mining is a technology or body of methods and techniques to analyze and discover processes from data[3] where data may be event logs referred as *history*, *audit trail* or a *transaction log*[4]. Similarly, Process Mining –or *Business Process Mining* as W. M. P. van der Aalst defines it[4]– can be described as the automatic construction of models which are explained in the event log[4].

1.1 *L* Life-cycle Model* Limitations and Proposed Extension

Given the considerable benefits Process Mining can bring to organizations, there is the challenge of taking Process Mining to a leading role in organizations that seek to improve business processes, operational efficiency or customer satisfaction. A number of methods for managing process mining

projects have been proposed, most notably W. M. P. van der Aalst's *L* life-cycle model*[5, p. 284-286] method, which provides a structure to frame Process Mining problems and match them with appropriate process mining techniques. Other proposals have been made such as the *PMD*[6] method or the PM^2[7] method. Given the availability of more than one Process Mining method or model, a discussion on the selection of the *L* model* is discussed in the *Background* chapter of this thesis.

The extension of a Process Mining model or method can be evidenced in Suriadi Suriadi Et. Al. use case where process behaviors in a large Australian insurance company [8, p. 461] are discussed. In such use case is stated that a waste of resources, time in this case, happened due to the lack of clear direction in order to answer relevant questions that may be raised by important stakeholders. This use case demonstrates the need to seek additional tools and techniques to properly structure the justification and elaborate a robust plan to conduct a Process Mining project.

While the *L* life-cycle model* is a good framework to structure and a guide Process Mining projects, it lacks the notion of *project charters* which is a Six Sigma's tool used within the *Define, Measure, Analyze, Improve, and Control (DMAIC)* model and that can contribute in positioning Process Mining projects from an effort supported by a small group of individuals or from a short-term project effort to a discipline that is embedded in an organization's culture. The L* *life-cycle* also misses guidance in the aspect of bridging Process Mining projects' goals with the objectives outlined by organizations' leadership and in a context of a Quality Management System or business improvement frameworks.

With the above limitations in place and considering the wide adoption of Six Sigma's *DMAIC* model and ISO 9001:2008 Quality Management System[9], the work in this thesis aims at extending the *L* model* with the methodology and QMS already mentioned. In regards to the business

improvement for excellence frameworks and considering the well established Baldrige Performance Excellence Program™[10] in the United States and the European Foundation for Quality Management Excellence Model™[11] in Europe, these frameworks have also been selected to extend the *L* life-cycle model*.

1.2 Thesis Outline

The second chapter of this thesis introduces the *L* life-cycle* model in every of its five stages and discusses alternative methods such as the Process Mining Project Methodology (PM^2) and the Process Diagnostics based in Process Mining method (PDM).

The third chapter of this thesis discusses how a Process Mining project using the *L** model can be justified for organizations that have adopted a Quality Management System or practice Business Improvement frameworks. In this chapter it is also detailed a Process Mining project via the *L** model for the Help Desk and SW Quality Assurance operations in an Italian IT Company where *stage 0: Plan* of the *L** model is extended with *DMAIC*'s *Define* phase by introducing a project charter, customer requirements as service level agreement (SLA) and building a top-level process model.

Fourth chapter elaborates in extending the *L** model *stage 1: Extract* phase by detailing a robust selection process of *key performance indicators* a.k.a. *KPIs* that can help project leaders and leadership in identifying the appropriate data to be extracted for the construction of a log. The chapter also works in extending the *Extract* stage with the *Measure* phase of the *DMAIC* model by detailing how data was extracted to arrive to a log file, the identification of measures, an overview of measures' statistics and business process and organizational mapping with Disco™[12] software. The

modeling of a detailed Help Desk business process with BPMN 2.0 language is also discussed.

The fifth chapter discusses the extension of the $L*$ model *stages 2 and 3: create control-flow model and connect event log and create integrated process model* in the context of how goals for Human Resources performance and competence, and the management of key processes and other resources can be identified in the basis of a Quality Management System and Business Improvement frameworks for a Process Mining project. This chapter also details the extension of the $L*$ model with *DMAIC*'s *Analyze* phase where causes of variation for conformance checking, time perspective and SLA metrics are identified with box plots, the ProM 6 XDotted chart and scatter plots, respectively.

Finally, chapter 6 discusses *stage 4: operational support* extension of the $L*$ model with audits' effectiveness by identifying the sources of errors that diminish the availability, reliability and suitability of audits and how Process Mining can contribute in diminishing the impact of these errors. The chapter also elaborates on extending the $L*$ model with *DMAIC*'s *Improve and Control* phases where a couple improvements for the Help Desk process are proposed and how these proposed improvements can be implemented and maintained.

CHAPTER 2

Background

W. M. P. van der Aalst's *L* life-cycle model*[5, p. 282-286], depicted in *Figure 2.1*, has been proposed to guide the execution of Process Mining projects, this model consists of five stages, *Plan and Justify, Extract, Create Control-Flow Model and Connect Event Log, Create Integrated Process Model,* and *Operational Support*. Each of these stages aims at accomplishing certain milestones in a Process Mining project where generally processes that are well structured can benefit from all of the *L* model* stages. For processes that are not well structured, these generally do not make use of the last stage of the model *Operational Support* because they remain within a process discovery phase.

2.1 Process Mining Methods

Several Process Mining Methods are available, notably the *PM² Process Mining Method*[7], and the *Process Diagnostic Method (PMD)*[6]. The PM² method is similar to the *L* life-cycle model* stages 0 and 1, *plan* and *extract*, in relation that the PM² method also has *planning* and *extraction* phases. Similarly the PMD method has a *log preparation* phase which could be matched to the stages mentioned for the other two methods.

The *L* model* and the PM^2 differ in the sense that the *L* model* creates the log during the *extract* phase while the PM^2 method achieves this task in the *Data Processing* stage.

L model* stages 2 and 3, *Create Control-Flow Model and Connect Event Log* and *Create Integrated Process Model*, are similar to PM^2's *Mining and Analysis* stages and PMD's *C, D* and *E* stages that perform *Control Flow Analysis, Performance Analysis* and *Role Analysis* activities.

Similarly, the PM^2 method deals with a *Evaluation* phase while the *L* model* details how *interpretation, intervention, adjustment and redesign* can reevaluate the *extract* stage of the *L** model where new questions and KPIs can be formulated and current questions and KPIs can be reevaluated. This feedback into the *extract* stage of the *L* model* is not a stage of its own but is clearly indicated by the model. The PMD method does not include a *process improvement & support* stage just as the PM^2 method does in its *evaluation* phase.

The *L* model* details in its stage 4 that *Operational Support* works mostly in assisting end users with prediction and recommendation features present in a given process. This could be interpreted as improvement efforts but this is not clearly defined in the *L* model*. In contrast, the PM^2 method does explicitly mentions that in its *process improvement & support* stage that *"The objective of the* ***process improvement & support*** *stage is to use the gained insights to modify the actual process execution"*[7, p. 305].

2.1.1 Process Mining Method Selection

In consideration of the findings detailed in the previous section, the *L* model* and the PM^2 method are better positioned as the ideal candidates to execute a Process Mining project that aims at improving a given business process. However, as it is detailed in the PM^2 paper *"L* covers more*

techniques, but was primarily designed for the analysis of structured processes and aims at discovering a single integrated process model."[7, p. 298] the *L* model* serves the purpose of working in structured processes with more techniques.

Given the broader scope of the *L* model* in terms of techniques and its focus in structured processes, this thesis discusses the extension of the *L* life-cycle model* since it better fits for organizations that have adopted a quality management system or practice a management initiative for excellence. Such organizations commonly, at least in theory, operate with structured processes. Similarly, the *L* model* has the notion of the KPI concept which is a widely adopted concept.

2.2 *L* Life-cycle Model* Stages

The initial stage of the model's ***Stage 0: Plan and Justify***[5, p. 284] states that in terms of justification for a Process Mining project, the organization sponsoring or executing the project could do this in one (or more) of the three types of proposed forms[5, p. 283]:

- A *data-driven project*
- A *question-driven project*
- A *goal-driven project*

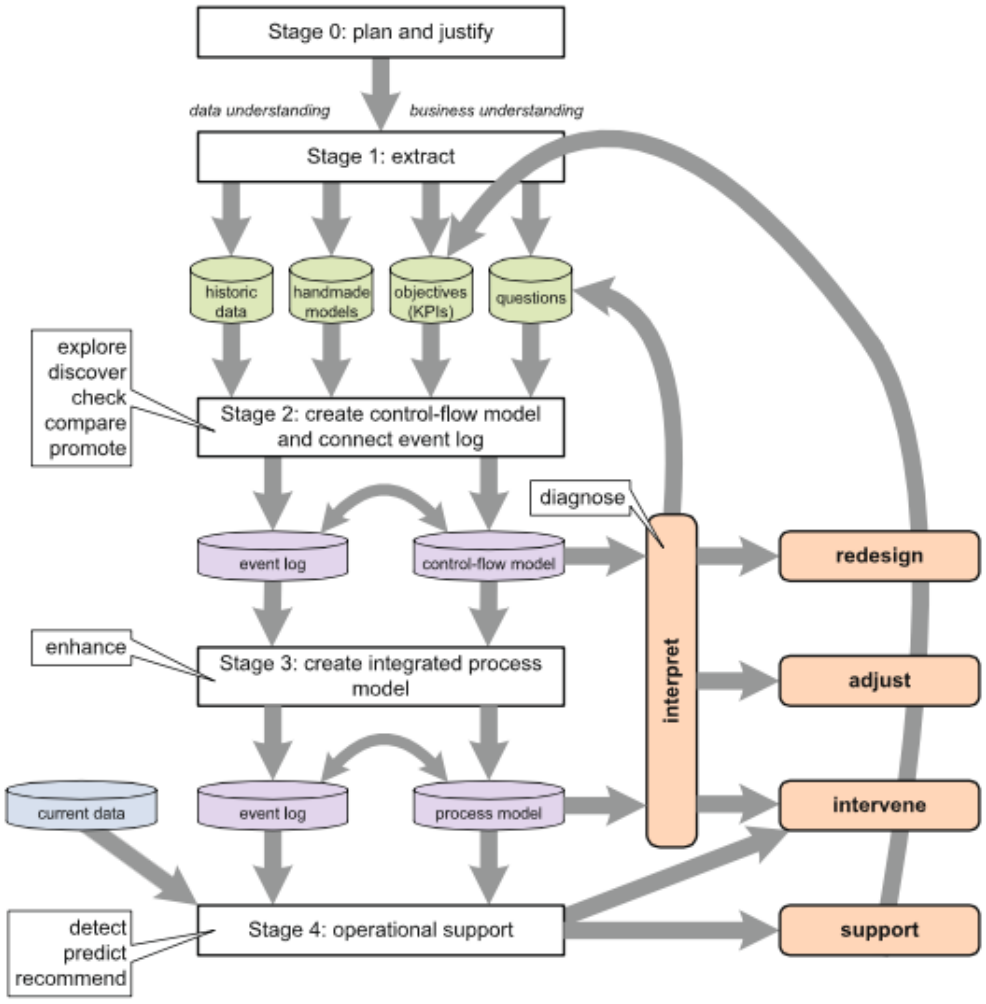
Once *Plan and Justify* stage is completed the model moves to the ***Stage 1: Extract*** which aims at extracting data from models, IT systems, experts in the project's domain and knowledge from management[5, p. 285]. A relevant piece of information in the *extract* phase is to build and event log, where in order to achieve this the log needs to have events ordered by time or time stamps and events need to be related to an unique case in the log.

Stage 2: Create Control-Flow Model and Connect Event Log is a step where the control flow is determined, in this context, there are two possible outcomes, a control flow is already modeled –say a documented process or de jure model– for which model synchronization known as *conformance checking* or other actions that attempt to answer questions related to the de jure model. If a process model is not in place, this stage consists in mapping a model with process discovery techniques[5, p. 285]. In the context of Process Mining, a process discovery technique is a function that, with the use of an algorithm and an event log, maps a process model such that *"the model is* ***representative*** *for the behavior seen in the event log"*[5, p. 125].

Stage 3: Create Integrated Process Model enhances the discovered models in previous stages by adding perspectives such as an organizational perspective or time perspective. This stage aims at having a comprehensive model that can serve several purposes[5, p. 286]. Processes that are not well structured are usually scrutinized up to this stage and are not further explored to the final stage of the *L^* model*. Processes that are well structured use the results of this stage as input for the final stage of the model. A structured process is one that can describe 80% of its behaviour contained within 20% of the process' instances[5, p. 148].

The last phase of the model, ***Stage 4: Operational support***, deals with *detect*, *predict* and *recommend* aspects of the analysis, in this stage is possible to predict the behavior of running process instances or cases and in this stage the goals of analyses are mostly oriented to users that are working in running cases[5, p. 286]. This stage of the *L^* model* is intended to only be used for structured processes.

FIGURE 2.1: *L* life-cycle Model*



The *L* life-cycle model* has been chosen due to its broad scope and [5, p. 284] characteristics where interpretation, redesign, adjust, intervene, and support feedback KPIs and questions.

2.3 Business Improvement for Excellence Frameworks

In this thesis two business improvement for excellence frameworks will be used to expand the *L* model*, these frameworks are introduced making mention on the organization that manage such frameworks.

2.3.1 European Foundation for Quality Management Excellence Model™

The European Foundation for Quality Management is a Brussels based non-for-profit membership foundation created in 1989 dedicated to increase the competitiveness of European businesses, the foundation's activities relate to training leadership and management on the use of the EFQM Excellence Model™ [11], assess organizations in using peer to peer review to realize the source of success, recognizing outstanding organizations, and sharing good practices that the Foundation recognizes through peer to peer assessments.

The EFQM Excellence Model™ is justified in the basis of determining whether organizations, regardless of their sector, structure, or maturity, need to establish an appropriate management framework[13, p. 2].

The model also *"assess where organizations stand in terms on the path to excellence and helps organizations in identifying their key strengths and potential gaps in relation to their stated Vision and Mission"*[13, p. 2]. The framework also provides a *"common vocabulary and way of thinking about the organization that facilitates the effective communication of ideas"*, additionally, the framework *"integrates existing and planned initiatives, removing duplication and identifying gaps"*, and *"provides a basic structure for the organisation's management system"*[13, p. 2].

In terms of concepts, the EFQM Excellence Model™ consists of seven fundamental concepts for excellence[13, p. 5-8]:

- *Adding Value for Customers*
- *Creating a Sustainable Future*
- *Developing Organisational Capability*
- *Harnessing Creativity & Innovation*

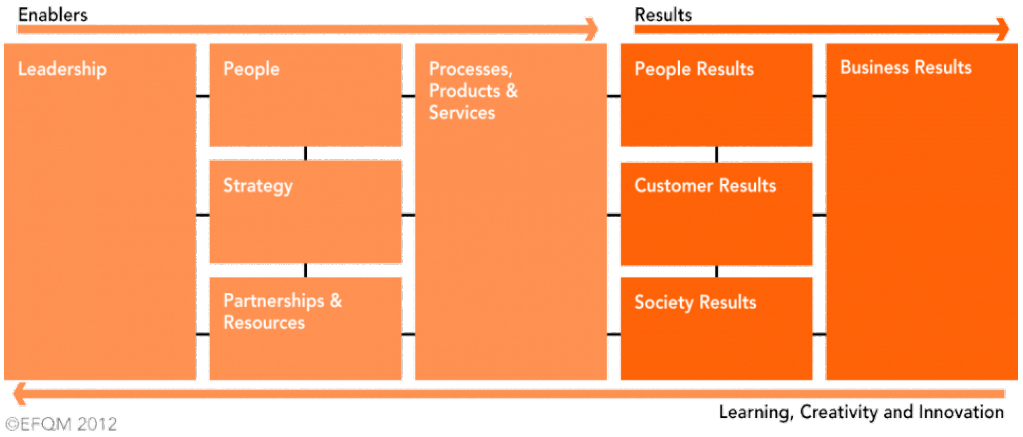
- *Leading with Vision, Inspiration & Integrity*
- *Managing with Agility*, and
- *Succeeding through the Talent of People*

Regarding criteria, the EFQM Excellence Model™ is made of *Enablers* and *Results* in the form of a *non-prescriptive* framework based in nine criteria[13, p. 10-20]:

- ***Enablers***
  1. *Leadership*
  2. *Strategy*
  3. *People*
  4. *Partnerships and Resources*, and
  5. *Processes, Products, and Services*
- ***Results***
  6. *Customer Results*
  7. *People Results*
  8. *Society Results*
  9. *Business Results*

The Model also details the need in identifying key processes through brainstorming, stakeholder interviewing, external services consulting or the Porter model[14]. The Processes, Products & Services enabler also details that the next step after identifying key process is the need to build the organization's top level processes models followed by modelling processes in more detail to a level that may include sub-processes that describe working levels, process' operations. So, re-engineering and step by step improvements can be made in order to enable efficiency and effectiveness improvements .

FIGURE 2.2: The EFQM Excellence Model™[1]



The model is a cause-and-effect diagram in which enablers drive results [11]. The *Processes, Products & Services* enabler can work in cooperation with the *L* life-cycle model* for Process Mining projects.

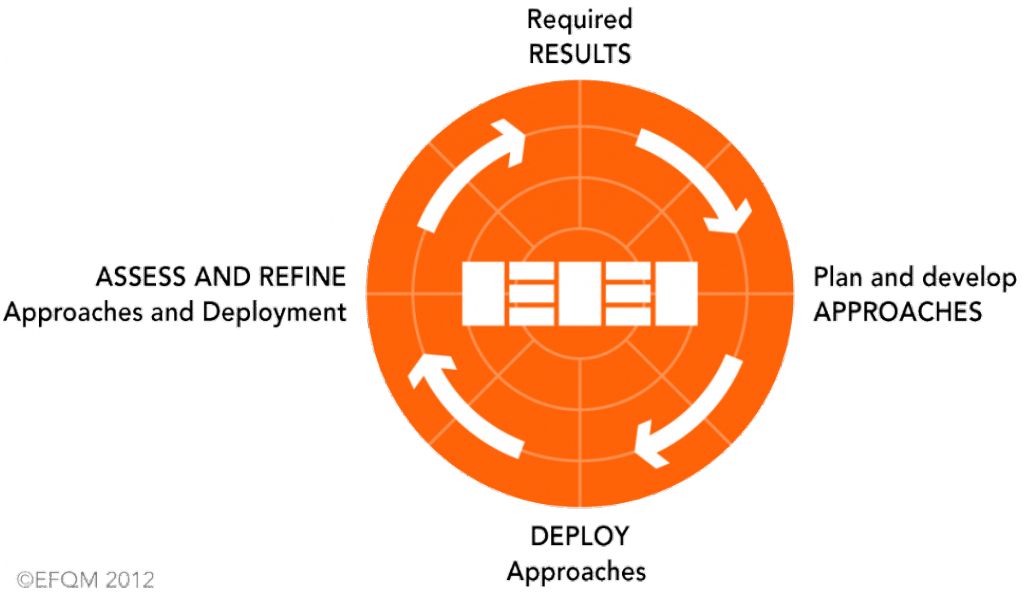
The Model also mentions the need to measure each part of the process in terms of cost, time and quality dimensions where a healthy balance of qualitative and quantitative measures should be in place. It is also imperative to manage and review process in such a way that [15]:

- *"Top leadership -the CEO or one of his or her direct reports- acts as the champion for the overall management and improvement of the process or processes in question".*
- There are clear roles and responsibilities to manage processes.
- There are cross-functional teams trained in process improvement.
- Recognition and rewards systems are in place where such systems are in line with processes' performance measurement.
- And that, there are appropriate measures in place.

Finally, the framework discusses the *RADAR* which is an assessment framework and a management tool that provides a structured approach for questioning the organization's performance[13, p. 22-25]. The tool states that an

[1]The EFQM Excellence Model™ is copyrighted material from the European Foundation for Quality Management, all rights reserved.

FIGURE 2.3: *RADAR* Logic™[2]



The RADAR logic is a dynamic assessment framework and management tool.

organization needs to *"determine the Results it is aiming to achieve as part of the strategy"*, *"plan and develop an integrated set of sound Approaches to deliver the required results for the present and the future"*, *"deploy approaches in a structured way to ensure implementation"*, and *"assess and refine the deployed approaches based in monitoring and analysis of the results achieved and ongoing learning activities"*.

In this context, the EFQM Excellence Model™ publication[13] provides a detailed matrix of the RADAR tool for the analysis of Enablers and Results where each element is broken down into a series of attributes[13, p. 23]. The cycle for the RADAR dynamic assessment is detailed in **Figure 2.3**.

[2]The RADAR Logic™ is copyrighted material from the European Foundation for Quality Management, all rights reserved.

2.3.2 Baldrige Excellence Framework™

The Baldrige Excellence Framework™ is administered by the Baldrige Performance Excellence Program a.k.a. *The Program*. The *Program* is based and managed by the National Institute of Standards and Technology which is an agency belonging to the United States Department of Commerce[16].

The *Program*, through its Baldrige Excellence Framework™ *"empowers organizations to reach its goals, improve results, and become more competitive"*[17, p. ii]. Similarly, the *Program* is dedicated in raising awareness on the importance of performance excellence, providing organizational assessment tools and criteria, as well as educating leaders in a wide range of organizations about implementing best practices.

The *Program* has instituted frameworks for excellence which have been tailored for three types of organizations[17]:

- *Baldrige Excellence Framework for Business and Nonprofit™*
- *Baldrige Excellence Framework for Education™*
- *Baldrige Excellence Framework for Health Care™*

The Baldrige Excellence Framework for Business and Nonprofit™ consists of three components [17, p. i], the *Criteria*, *core values and concepts*, and *scoring guidelines*. The porpoise of the framework is helping organizations –regardless of size, sector, or industry– answering three questions: *"Is the organization doing as well as it could? How is this known? What and how should the organization improve or change?"*[17, p. ii]. Therefore, the *Criteria* contained in the aforementioned framework helps organizations in responding such questions.

The *Criteria* is then divided into further questions which are made of *"six interrelated process categories and a results category"*[17, p. ii]. These categories are[17, p. ii]:

1 *Leadership*

2 *Strategy*

3 *Customers*

4 *Measurement, analysis, and knowledge management*, and

5 *Workforce*

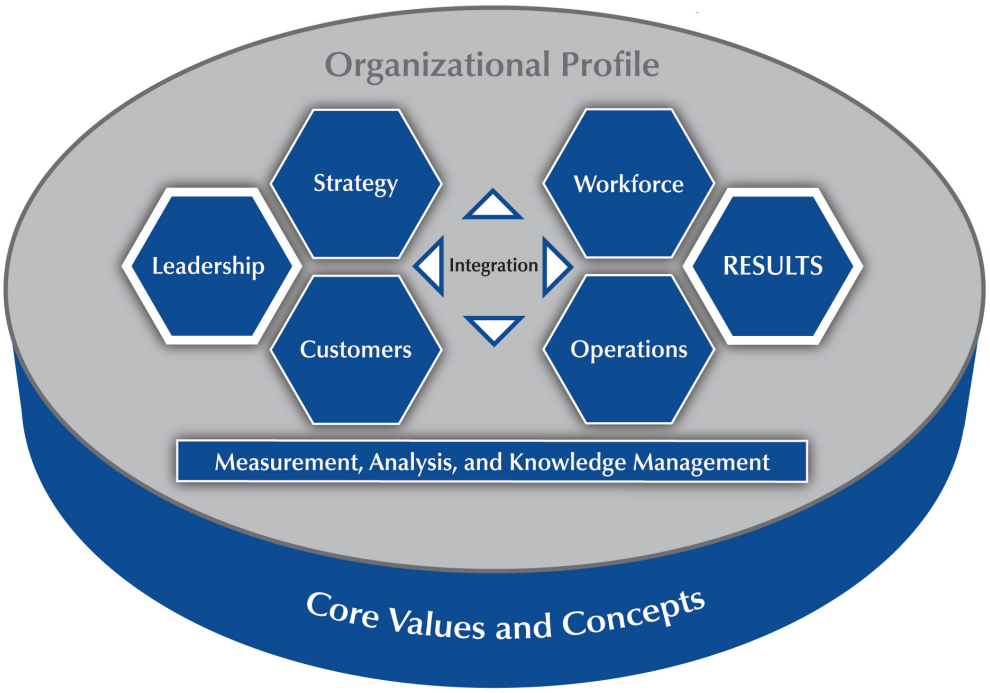
6 *Operations*,

7 *Results*

In this context the Baldrige Criteria™ can be modeled as depicted in **Figure 2.4**. One of the key aspects for The Baldrige Excellence Framework is the concept of *Measurement, Analysis and Knowledge Management* which emphasizes that management[18] must focus on performance excellence for the entire organization. These key elements could enable the justification for a Process Mining project with the *L* model*.

In this thesis, we will discuss how the questions formulated in the aforementioned categories within *Criteria*, a.k.a. the ***Baldrige Criteria***, can be used to expand the *L* Life-cycle Model*. A detailed questionnaire for each one of the categories within the *Baldrige Criteria* for the Baldrige Excellence Framework for Business and Nonprofit can be acquired at the *Program*'s website.

FIGURE 2.4: The Baldrige Excellence Framework for Business and Nonprofit[3]



From Baldrige Performance Excellence Program. 2015. *2015–2016 Baldrige Excellence Framework: A Systems Approach to Improving Your Organization's Performance.* Gaithersburg, MD: U.S. Department of Commerce, National Institute of Standards and Technology. http://www.nist.gov/baldrige.

[3]This figure is used with permission of the Baldrige Performance Excellence Program. 2015. 2015-2016 Framework for Performance Excellence. A systems approach to improving your organizations performance. Gaithersburg, MD: U.S. Department of Commerce, National Institute of Standards and Technology. Obtain a copy of the full Framework at `www.nist.gov/baldrige/publications/criteria.cfm`

2.4 ISO 9001:2008 Quality Management System

As a global organization for standardization, ISO has been leading the standardization of a quality management system with ISO 9001, the ISO 9001:2008 Quality Management System (QMS) is the latest version of such standard which is widely adopted among many organizations, public and private. According to the ISO Survey [19], there were about 1.1 million ISO 9001 certificates issued in 2013. While there are other ISO certifications such as ISO 14001 or ISO/TS 16949 for environmental management and the automotive industry respectively, among other, we will focus in the ISO 9001:2008 QMS as this is the most widely applicable one.

The ISO 9001:2008 QMS is made of several clauses and subclauses, which are listed as follows:

Clause 4 *Quality management system*

- **4.1** General requirements
- **4.2** Documentation requirements

Clause 5 *Management responsibility*

- **5.1** Management commitment
- **5.2** Customer focus
- **5.3** Quality policy
- **5.4** Planning
- **5.5** Responsibility, authority, and communication
- **5.6** Management review

Clause 6 *Resource management*

- **6.1** Provision of resources
- **6.2** Human resources
- **6.3** Infrastructure
- **6.4** Work environment

Clause 7 *Product realization*

- **7.1** Planning of product realization
- **7.2** Customer-related processes
- **7.3** Design and development
- **7.4** Purchasing
- **7.5** Production and service provision
- **7.6** Control of monitoring and measuring equipment

Clause 8 *Measurement, analysis and improvement*

- **8.1** General
- **8.2** Monitoring and measurement
- **8.3** Control of nonconforming product

8.4 Analysis of data
8.5 Improvement

In subsequent chapters it will be discussed how some of the requirements set forth by ISO 9001 QMS can be used to expand the *L* model*.

2.5 Six Sigma's DMAIC Model

According to Pyzdek and Keller Six Sigma handbook *"Six Sigma is a rigorous, focused, and highly effective implementation of proven quality principles and techniques"*[20, p. 3] On of these techniques is *"a simple performance improvement model known as Define-Measure-Analyze-Improve-Control, or DMAIC"*. The *DMAIC* model is briefly described as follows[20, p. 4]:

D Define the goals of the improvement activity.

M Measure the existing system.

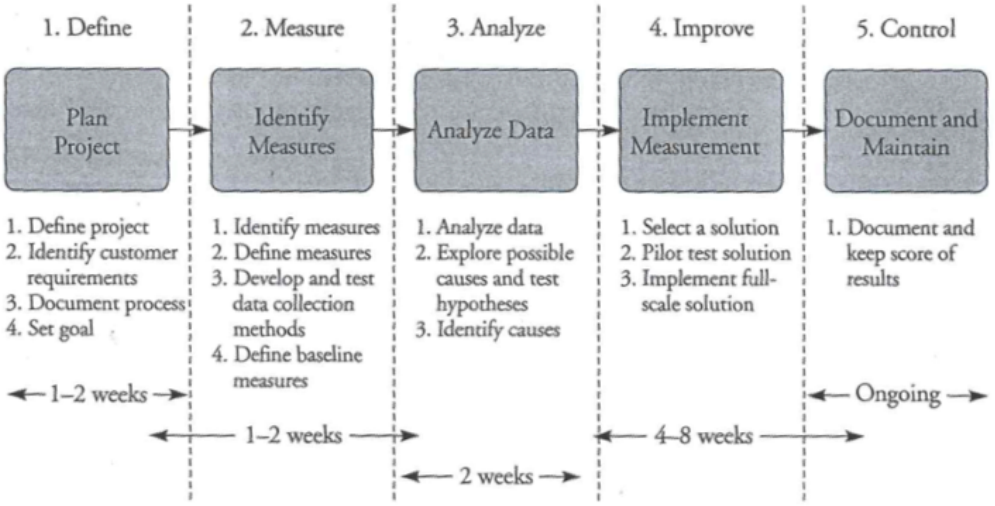
A Analyze the system to identify ways to eliminate the gap between the current performance of the system or process and the desired goal.

I Improve the system.

C Control the new system.

In consideration that the *L* model* will be extended with Six Sigma's *DMAIC* model, the project will use the model proposed by Harmon's *Business Process change*[21] publication and detailed in **Figure 2.5**. This model shows the typical activities that are to be conducted during a Six Sigma project for business process improvement, while not all the activities indicated by the model will be performed during the work of this thesis, the model does provide valuable guidance for the project team to structure goals and progress indicators.

FIGURE 2.5: Six Sigma Project Overview



Harmon's Proposed overview of a Six Sigma project[21, p. 325] details the typical activities to be conducted for business process improvement in the context of a Six Sigma project.

CHAPTER 3

Extending L* *Life-cycle Model* Stage 0: Plan and Justify

This chapter discusses how a Process Mining project following the *L* life-cycle model* can be extended using ISO 9001:2008 Quality Management System and Business Improvement frameworks. The Help Desk Process Mining project will also be introduced and it will be explained in detail the definition of the project, the identification of customer requirements and the modeling of top-level process model using BPMN 2.0 language.

While the author of the L* *model* clearly articulates what could be the three most common process mining projects –as mentioned in 2.2–, it is important to mention that Process Mining projects do not need to be necessarily confined to these three types of projects. In fact, many organizations could justify a Process Mining Project using the *L* model* because ISO 9001:2008[9] Quality Management System (QMS) and Business Improvement frameworks dictate the need for organizations to map and document key processes as well as understanding how resources interact among each other, therefore the motivation for organization to use Process Mining to achieve these goals. Concerning Six Sigma's *DMAIC* model and its *Define* phase, this can be an asset for the planning of Process Mining projects using

the *L* model*, therefore extending the model where objectives are identified and communicated within the stakeholders involved in the project.

3.1 Justifying a Process Mining Project with ISO 9001:2008 QMS

Given the need for organizations that have adopted the ISO 9001 QMS in documenting, controlling and improving processes, this section will discuss how the ISO 9001 QMS can be used to expand *L* model justify* stage.

The ISO 9001:2008 standard states, in its introduction section, that:

- *"An organization's design and implementation of a quality management system are influenced by, among other things, the process it employs"*[9, p. v].
- In section 0.2 for *process approach* [9, p. v-vi] is also stated that *"the standard promotes the adoption of a process approach during the development, implementation and improvement of a quality management system in order to improve customer satisfaction"*. Furthermore, the requirement mentions that *"for an organization to work efficiently and effectively, it must manage numerous activities that are related between each other"*[9, p. v].
- In terms of activities, the standard also in its *0.2 Process approach* section, states that *"an activity or a set of activities that utilize resources and where such activities are managed with the objective that the input elements are transformed in results can be considered a process"*[9, p. v]. Finally, it is also stated that, *"frequently, the result of a process is the input of an other process or the next process"*[9, p. v].
- Additionally, in section 0.2, it is stated that *"the application of a process oriented system within an organization together with the identification and interaction of such processes, as well as its management,*

> *in order to produce the desired output can be determined a process approach"*[9, p. v].

In the aforementioned section, it is finally stated that some of the important points of a process based approach are:

- A need to consider processes that add value
- Obtain results and effectiveness from a process
- The continual improvement of a process based in objective measurement

In Process Mining language, a *de jure model* can be related to a *documented process* in a QMS context. De jure models are documented processes which are a requirement of the ISO 9001:2008 QMS, documented processes, de jure models, and normative models mean the same thing and will be used alternatively during this thesis in the form of business process models in BPMN 2.0 language, in the form of Petri nets or other notations.

Given the above points detailed in ISO 9001:2008 *0.1 General and 0.2 Process Approach*, the QMS details a model of a *Process Based Quality Management System*. The model depicted in **Figure 3.1** clearly indicates that a fundamental part of the cycle is *the product (or service) realization* and the *measurement, analysis and improvement of the quality management system*. Given these requirements in the QMS model and aforementioned clauses, the justification of a Process Mining project can be made based on these requirements. Since some of the Process Mining techniques are related to the discovery of a *real* business processes and the comparison of these to de jure models, it is then in the interest of organizations to leverage from the benefits of Process Mining to achieve the goals dictated in the QMS.

Clauses 7 and 8 state that *"an organization shall plan and develop processes for the product, or service, realization"*[9, p. 7]. In particular, in section 7.1

the requirements indicate that during the planning for product realization the organization needs to *"establish process and documents"* and *"have records"* in order to provide evidence for product or service realization and meeting product or service requirements.

In these terms, it is worth mentioning that one of the Process Mining approaches is to establish a de jure model, in order to perform business processes auditing [5, p. 191]. Therefore, when an organization is designing its processes it could use the *conformance checking* a.k.a. *synchronization* Process Mining approach to ensure that processes are not only documented, but that they are actually being followed.

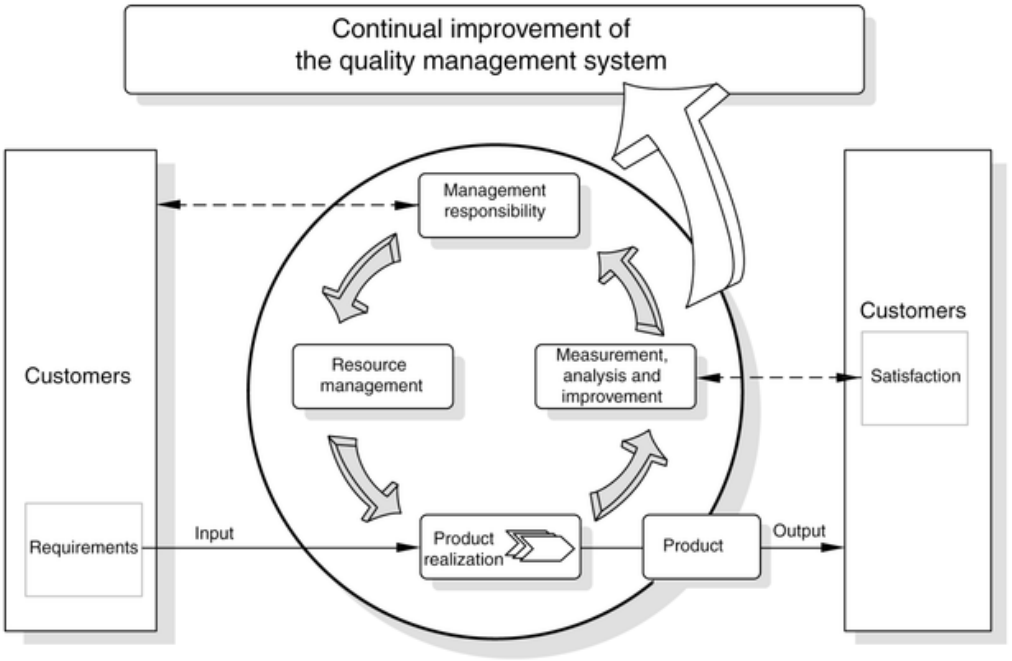
Conformance checking is defined as *"comparing observed behavior with modeled behavior"*[5, p. 192]. Conformance checking measures can be defined in three categories[5, p. 192]:

- The overall conformance of the model and the entire log.
- Local conformance diagnostics for nodes in the model.
- And, Conformance checking measure for specific cases.

To conclude this section, an additional clause in the ISO 9001:2008 QMS is the *Measurement, analysis and improvement* which [9, p. 12] states that organizations:

- *"Shall plan and implement the monitoring, measurement, analysis and improvement processes needed"*[9, p. 12]
- In section 8.2.3 the standard also states that *"the organization shall apply methods to monitor and measure the quality management system"*[9, p. 12] where these methods should be able to *"demonstrate the ability of processes to achieve the planned results"*[9, p. 12].

FIGURE 3.1: Model of a Process-based Quality Management System[1]



During the product realization and measurement, analysis and improvement phases[9, p. vi], Process Mining projects can be justified in order to aid organizations for the execution of these stages.

- Additionally, monitoring and measurement of product, section 8.2.4, details that *"characteristics of the product shall be monitored and measured in order to verify that product requirements and have been met"*[9, p. 12].

It is important to mention that when the word *product* is referred, this means that the product could either be a tangible or non tangible output. In the context of ISO 9001:2008 QMS in its clauses 7 and 8 as detailed above, it is clear that a Process Mining project can be well justified, because the standard requires that processes must be designed and these processes shall also be monitored and measured, Process Mining can then have significant synergy with the standard by aiding organizations in understanding whether the established processes to realize products or services are being executed properly and other several aspects of such processes are on target.

[1]The *Model of a process-based quality management system* is copyrighted material from the International Organization for Standardization (ISO), all rights reserved.

3.2 Process Mining, the EFQM Excellence Model™ and Baldrige Criteria™

Some organizations seek business improvement outside the adoption of a quality management system, so an alternative for such practice is the adoption of the Baldrige Criteria™ or the EFQM Excellence Model™. In this context, Process Mining can work with these frameworks in order to assist organizations in achieving their goals.

3.2.1 *L* Model* and the Baldrige Criteria™

Since the Baldrige criteria decrees the the leadership in organizations must manage their process by fact, the criteria is well positioned to be used as an extension of the *L* model*. In particular because Process Mining tools and techniques rely in greater part by data persisted in information systems, an asset for leaders to be able to approach processes based in facts.

Therefore, the criterion is relevant since it highlights the fact that management must be aware of facts and manage based on them, this is stated as one of the core values of the Baldrige Criteria™ as *Management by fact* [22]. As David Garvin mentions [23]: *"To win, companies must have customer-oriented quality programs that are led by senior management, a high level of employee involvement, an understanding of internal processes, and* ***management by fact*** *rather than by instinct or feel."* This quote very well matches one of the points of W. M. P. van der Aalst's Process Mining book where it is stated that *"Using recent breakthroughs in process mining, we will show that it is possible to simplify and unify the analysis of business processes based on facts"*[5, p. 23].

As such, where management has to focus and take decisions based on facts, Process Mining through the *L* model* can certainly have an important effect in achieving such goals because one of the Process Mining techniques is discovering process models from logs –persisted data– and compare these discovered models to normative models. Normative models that an organization may have determined beforehand and that are an important element in managing and improving processes. In consequence, the *L* model* can have a significant and positive impact in helping management in managing by fact rather than studying processes solely via interviews or *instinct*.

3.2.2 Process Mining *L* Model* and The European Foundation for Quality Management Excellence Model™

The EFQM Excellence Model™ being a cause-and-effect relations model, depicted in **Figure 2.2**, has *enablers* and *results*, one of the key enablers is the *Processes, Products & Services* component[11]. The Processes, Products & Services enabler, mentions that *"processes are a sequence of activities that add value through output from several inputs where such inputs could be the output(s) of other processes"*[15]. Additionally, the Model specifies that each activity or process step has four basic elements to address:

- Suppliers
- Customers
- Performance Indicators, and
- Control Loops

Considering that the EFQM Excellence Model™ Processes, Products & Services enabler criteria, this can be used to expand Process Mining projects with the *L* model* in its *justification* and *extractions* phases. The *L* model*

stage 1 emphasizes in the need to use data through an extraction process based in questions and objectives such as *key performance indicators*. So, processes can be mapped and be compared to a normative model as discussed in *L* model* stage 2, as well as recommending or predicting the improvement or behaviour of a process, respectively in the stage 4 of the model. In consideration of the Processes, Products & Services enabler criteria from the EFQM Excellence Model™ where it is mentioned that *"5.a Processes are designed and* ***managed*** *to optimize stakeholder value"*[13, p. 16] and *"Use data on the current performance and capabilities of their processes, as well as appropriate benchmarks, to drive improvement ..."*[13, p. 16], such statements can be qualified as a justification statement in order for organizations to drive Process Mining projects with the *L* model*. As mentioned in 3.2.1, one of the Process Mining techniques is discovering processes models from logs in order to compare these discovered processes to normative models. So, the EFQM Excellence Model™ could service the *L* model justify*, effectively expanding it.

3.3 Extending Process Mining *L* Life-cycle Model* Planning Stage with Six Sigma's *Define* Phase

Process Mining, Lean, and Six Sigma's *DMAIC* model, working in synergy, is an idea that has been briefly discussed in W. M. P. van der Aalst's Process Mining book. In this publication is mentioned that Six Sigma, in terms of tools and techniques, intends to improve the quality of processes[5, p. 22] based in the *DMAIC* approach detailed in **section A.1** appendix.

With this in mind and to illustrate how the *Define* phase of the *DMAIC* model can be applied in the *L* model*, we will use as an example the Help Desk process of an Italian IT company were real data was used. This section of the chapter then, will discuss how the *L* model plan* stage can in practice be expanded with the aforementioned phase of the *DMAIC* model.

An important aspect of any project is to identify the benefits, stakeholders and specific KPIs that the project is going to consider. In this context, a tool used during the *Define* phase is the drafting and approval of a *project charter*. The Project Charter consists of:

- Defining the scope, objectives, and schedule of the project
- Defining the process that the project will be focusing and its stakeholders
- Selecting the team members
- Obtaining authorization from sponsor/top management

A drafted project charter for the project performed in the Italian IT Company is detailed in **section A.2** appendix. Due to Company's request, the data, stakeholders, and other relevant and sensitive information has been anonymized as applicable.

While the project does not consider achieving costs savings, the project does discuss in what measure the de jure model –once this has been composed– is being executed by the stakeholders, and what are the interactions between the Help Desk Process and SW Quality Assurance stakeholders in the process. The project also has a goal to understand how the different software releases for one of the most important company's product is impacting the amount of incidents logged in the Help Desk process. In summary, the project charter drafts the following goals for the project which were completed in a four month span:

- Build a de jure model taking in consideration event logs as well as stakeholders input via interviews to compare the de jure model and event logs to understand in what measure the de jure model is being followed.

- Stakeholder interaction in Help Desk process and Bugzilla process as well as stakeholders interaction in these two processes.
- Measure in what degree the service level agreement (SLA) metrics are being followed.
- Analyze what are the different data attributes related to help desk incidents and how these impact the SLA.
- Measure and analyze how the different software releases are influencing the number of Help Desk incidents related to potential SW anomalies issues.
- Based in the above analyzes, draft recommendations, if any, and deploy improvement and control measures where applicable.

The goals stated above are well within Harmon's Six Sigma approach for process improvement where it is mentioned that many Six Sigma projects begin by helping leadership develop a process architecture where the project typically lasts from one month to six months[21, p. 322-323].

Similarly, Harmon's work discusses about the team formation for the project, in our case, the project has focused in selecting team members for the project who have expertise in the process being evaluated[21, p. 324]. Another note form Harmon's work is the need to identify customer requirements[21, p. 325], in this context, our project will heavily focus in measuring the SLA durations and the analysis of such measurements. In **Table 3.1**, a service level agreement poster available for customers indicates the maximum amount of time the company should take to resolve incidents. This table is important since it is going to be a parameter that will be used to determine whether the agreements may be met consistently for the several incidents logged by customers via the Help Desk operation. As well, Table 3.1 details what are the three levels of service –*Basic, Advanced, Full Service*– available for costumers to purchase and how incidents can be categorized

TABLE 3.1: Service Level Agreement (SLA)

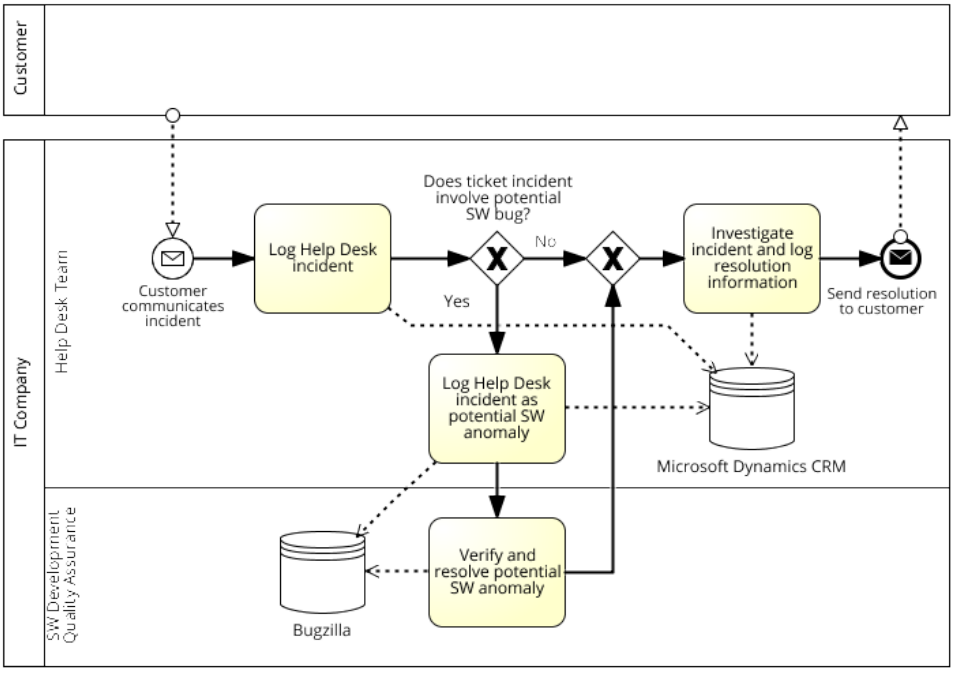
Severity	**Status**	Service level		
		Basic	*Advanced*	*Full Service*
Level 1 red code	Assign Gravity	4 hours	2 hours	1 hour
	Taking charge	12 hours	8 hours	4 hours
	Resolution	24 hours	16 hours	8 hours
Level 2 yellow code	Assign Gravity	4 hours	2 hours	1 hour
	Taking charge	32 hours	16 hours	8 hours
	Resolution	12 days	6 days	3 days
Level 3 white code	Assign Gravity	4 hours	2 hours	1 hour
	Taking charge	10 days	5 days	2 days
	Resolution	New Release	40 days	20 days

in three different severity levels *Level 1 red code, Level 2 yellow code, or Level 3 white code.*

A *Level 1 red code* incident is the most urgent one as it denotes the complete impairment of an IT system where all of its users are unable to access it or use it. The *Level 2 yellow code* denotes an incident where one or more users are unable to use one or more of the functionalists offered by the IT system in question. A *Level 3 white code* is an incident where a secondary functionality of the IT system is preventing any user to perform the usual tasks. Table 3.1 then summarizes the maximum amount of time it should take for activities to be processed in terms of incident's severity and service level.

An other important activity in the *Define* phase is to identify a top-level process map, in this project, this goal was achieved by discussing with stakeholders what are the main activities performed by them and what data these activities generate. Because our project focuses on two processes, Microsoft Dynamics CRM and Bugzilla, it was concluded that these two processes persist data in two data bases. Data from these two entities will be extracted to create a .XES log for the project. **Figure 3.3** details our top-level process map where it can be deduced that two main processes interact, Microsoft Dynamics CRM –a.k.a. CRM– process and Bugzilla process.

FIGURE 3.2: Top-level process map for CRM and Bugzilla for a Help Desk Operation



Bugzilla is *"a Defect Tracking System" or "Bug-Tracking System"* which allows *"individuals or groups of developers to keep track of outstanding bugs in their product effectively"*[24].

The top-level process starts when a customer communicates an incident, this is then logged in the data base for the CRM process, an incident may or may not lead to a potential SW anomaly incident in which a Bugzilla ticket may be created. If the incident involve a potential SW anomaly, this is verified and resolved by the SW Quality Assurance team who executed the Bugzilla process to manage the SW anomaly, a set of activities that are persisted in the Bugzilla process data base.

In summary, the top-level process map in Figure 3.3 shows that after a customer reports an incident, this could be related or not to a potential SW anomaly. As the incidents are processed, data is stored in CRM and Bugzilla data bases as applicable.

CHAPTER 4

Extending *L* Life-cycle Model* Stage 1: Extract

In the previous chapter we discussed how the *L* model* can be extended in terms of justifying a Process Mining project for organizations that have adopted the ISO 9001:2008 QMS, the Baldrige Criteria^{TM} or the EFQM Excellence Model^{TM}. We also mentioned how a Process Mining Project can be properly planned with *DMAIC*'s *Define* phase, effectively extending the *L* model* by aiding in identifying the goals for the project, the team members, identifying project's opportunities, elaborating on the business impact by conducting the project, and made a list of process indicators in terms of deadlines and time line planning.

Therefore, the next stage in the *L* model* is the extraction of data, for this, it is paramount to clearly identify what type of data is available and what data is to be extracted. In order to properly select the right data to be extracted is important to clearly identify what are the KPIs that will be evaluating the process. As well, in this stage it is necessary to conduct qualitative research to gather information so a hand made model or a de jure model can be made.

In regards to ISO 9001:2008 QMS and Baldrige Criteria™, as well as the EFQM Excellence Model™ frameworks, these will be discussed in order to be able to expand the *Stage 1: extract* part of the *L* model* in terms of how a Process Mining Project can benefit from these frameworks to identify KPIs, gather or generate de jure models.

4.1 *L* Life-cycle Model* Data Extraction: KPIs Selection and De Jure Models

L life-cycle model Extract* stage consists in gathering documented process models, project objectives, data, and other relevant questions from IT systems, process owners, management and domain experts, as applicable. In this context, such process models may be documented if an organization practices a quality management system or business excellence initiatives frameworks, a set of relevant questions could be in the form of critical *KPIs* that the organization can select based in the capabilities that the Process Mining project could deliver. Other important questions could be formulated by the leadership that could, for example, be taken from criteria examinations. Given that the *L* model* in its *extract* stage requires the formulation of the aforementioned criteria and because the KPI concept is a well established practice in the context of a quality management system and business excellence initiatives, this chapter discusses how the *Extract* stage of the model can be extended with a robust selection of performance indicators.

4.1.1 ISO 9001:2008 KPI Selection Criteria

A KPI is not to be confused with a target, to make a clear definition of a KPI this has to be expressed as a ratio or percentage so this is tracked

over time to be able to observe trends in performance. An example of a KPI is *"...the rate at which shipped items are returned..."*[25]. Additionally, targets can be in the performance data so it can be induced whether the given KPI is under or over performing [26, p. 53].

In Ian Rosam and Rob Peddle publication is detailed that the ISO 9001:2008 standard requires that processes' and products' performance are to be measured against targets where this and other information is to be analyzed to be able to come with opportunities for improvement [26, p. 53]. Similarly, the ISO 9001:2008 standard in its clause number 8 *Measurement, Analysis and Improvement* details that organizations monitor information related to customer perception in order to be able to determine if the organization is meeting customer requirements[9, p. 12]. In this context, identifying the right KPIs, as the standard requires, not only have to be related to processes and products but must also to be linked to business objectives [26, p. 53].

Considering that *L* model extract* stage consists of acquiring *event data, models, objectives, and questions*[5], these activities then need to take in consideration whether they will meet the objective of helping the project team in evaluating the processes where the data is being extracted. Such evaluation can take several forms where one of them can be a set of KPIs. In the light of this, a good understanding, design and selection of KPIs before the data extraction takes place can benefit a project that uses the *L* model*, consequently extending the *extract* stage of the model.

In order to determine if the organization has the right KPIs, it is important to take in consideration the following:

- KPIs must be suitable for the organization that has control on the process that is being measured, as it is not desirable that an organization or process is measured with a KPI which the stakeholders can not control.

- Is important to have the right amount of KPIs, too many KPIs may cause the organization or those working in a process to lose objectivity.
- ISO 9001:2008 QMS clause 6 in its sub-clause 6.2.2 *Competence, training and awareness* specifies that *"it must be ensured that personnel is aware of the relevance and importance of their activities and how they contribute to the achievement of the quality objectives"*[9, p. 6]. So resources (people) working in a process must be aware how their activities influence the KPIs.
- It is preferable that a resource working in a process has not more than 3 KPIs that are monitored simultaneously, preferably 2.
- In Rosam and Peddle publication [26, p. 56] a reference on whether system KPIs are suitable is available in **section A.3** appendix.

In **Figure 4.1** a continuous improvement cycle for system and process KPIs is depicted, here is important to relate these continuous improvement cycles to the *L* model* which details how KPIs can be redesigned, adjusted, intervened, or supported based in input from the model's stages 2, 3 and 4. As a conclusion, identifying the right amount of KPIs and the right KPIs where stakeholders are aware how their actions influence the objectives is a paramount activity before the extraction of data takes place. Since organizations may have vast amount of data, it is therefore wise to first design and select the right KPIs so these can be monitored by extracting the right data.

FIGURE 4.1: Rosam and Peddle System and Process KPI Continuous Improvement Cycle [26, p. 55, 57]

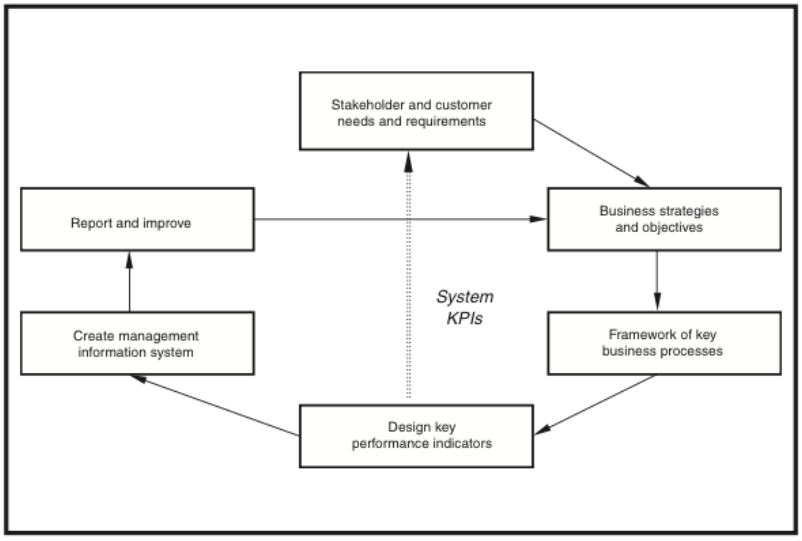


Figure 7.1 System (strategic) KPIs and the continuous improvement cycle at an organizational level

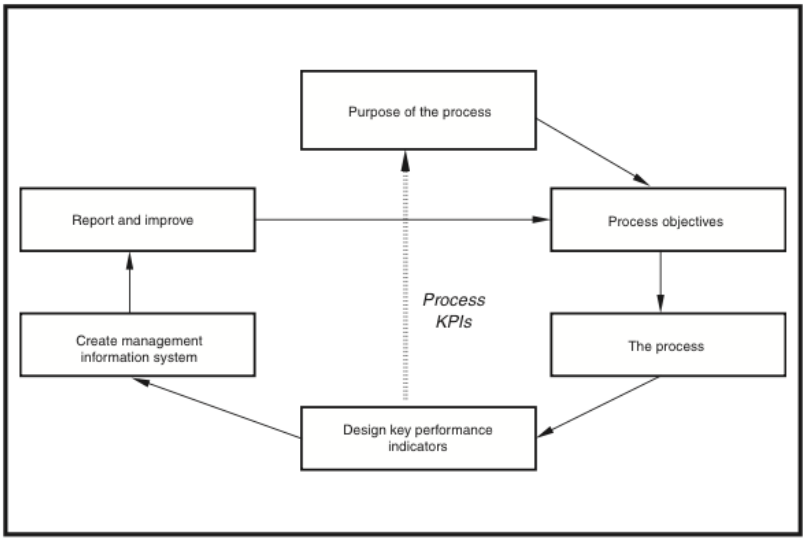


Figure 7.2 Process (tactical) KPIs and the continuous improvement cycle at a process level

4.1.2 EFQM Excellence Model™ and Baldrige Criteria™ KPIs

While the Baldrige Criteria for Performance Excellence™ and the EFQM Excellence Model™ do not explicitly mention the KPI concept, the need to be able to identify key performance measures or indicators is widely mentioned in these frameworks, due to the nature of the *L* model* which aims at working around several aspects of processes, this subsection will focus in business improvements for excellence criteria that discuss processes.

The EFQM Excellence Model™ in its fifth criteria related to *Processes, Products & Services* details in sub criteria *5d* that *"Excellent organizations compare their performance with relevant benchmarks and learn from their strengths and opportunities fro improvement in order to maximize the value generated for customers."*[13, p. 16] This criteria could be a starting point for the leadership to asses the need on whether a Process Mining project could be the right tool to map processes in two organizations, to later on compare aspects of these processes such as conformance *conformance checking* where a critical KPI could be the number of deviations in the model over the number of activities executed in a given time frame.

Similarly, criteria *5a* states that *"Excellent organizations develop a meaningful mix of process performance indicators and related outcome measures, enabling the review of the efficiency and effectiveness of the key processes and their contributions towards the strategic goal."*[13, p. 16] This criteria essentially dictates the need to implement key performance indicators for key process across the organization, if the organization looking to adopt the criteria has the means to support a Process Mining project then the selected KPIs can be integrated in the project's *L* model extract* stage.

In terms of the Baldrige Criteria for™, criteria 2.2 *Strategy Implementation* details in section 5 *Performance Measures* that organization shall

ask whether *"key performance measures or indicators are used to track the achievement and effectiveness of the action plans"*[17, p. 12]. This Baldrige Criteria™ can also expand the *Extract* stage of the *L* model* by integrating as part of the Process Mining project KPIs formulated based in such Baldrige Criteria™. In this context, the Baldrige Criteria™ also elaborates on the need to question whether the organization is formulating projections in terms of key performance indicators, on this, Baldrige Criteria™ (also in *Strategy Implementation*) details in section 6 *Performance Projections* if organizations *"have performance projections for short and longer term planning horizons"*. This criteria could also be integrated as part of the *Extract* phase of the model for Process Mining projects that are planned to be executed in the future.

It is then evident that business excellence initiatives can bring considerable value when a Process Mining project within the *extract* phase of the *L** model, in particular, key performance indicators often times are formulated by the leadership for which such formulation could consider the possibility to measure such indicators within the discipline of Process Mining.

4.2 Extending Process Mining *L* Life-cycle Model* Extract Stage with Six Sigma's *Measure* Phase

In the previous chapter we discussed how our Process Mining project based in the *L* model* has been extended with *DMAIC*'s *Define* phase in order to properly draft the project's charter so scope, objectives and schedules are identified and communicated as well as creating a team that will be working in the project. In the Define phase we also constructed a top-level process map to understand what processes are generating data that can be extracted for our Process Mining Project, and identified numeric parameters such as SLA.

With this in mind, this subsection will focus in the next stage of Six Sigma's DMAIC model, the *Measure* phase. In our project charter, **section A.2** appendix, we defined that some of the project's goals are:

- Build a de jure model taking in consideration event logs as well as stakeholders input via interviews.
- Understand how the stakeholders are interacting in the Help Desk and Bugzilla processes.
- Understand in what degree SLA metrics are being met.
- Understand what are the different data attributes related to help desk incidents and how these impact the SLA.
- Understand how the different software releases are influencing the number of Help Desk incidents related to potential SW anomalies issues.

To meet the goals outlined above, we need to find out what data can help the project team as well as building a business process model based in interviews and using Process Mining software. To achieve these goals, we will discuss how the data extraction task took place and a brief explanation of the log file will be provided, this will be discussed in the following subsection. In subsection 4.2.2 we will discuss the modeling of a business process with stakeholder interviews and Process Mining software.

4.2.1 Data Extraction and Log File Statistics Overview

We already discussed that data is stored in Microsoft Dynamics CRM and Bugzilla data bases, so the next natural step is to identify what tables from these systems are the ones to be used so a log file can be constructed. Because the SLA is an important aspect in our project, it is imperative that our data includes time stamps so the amount of time activities take

FIGURE 4.2: Data Extraction Model

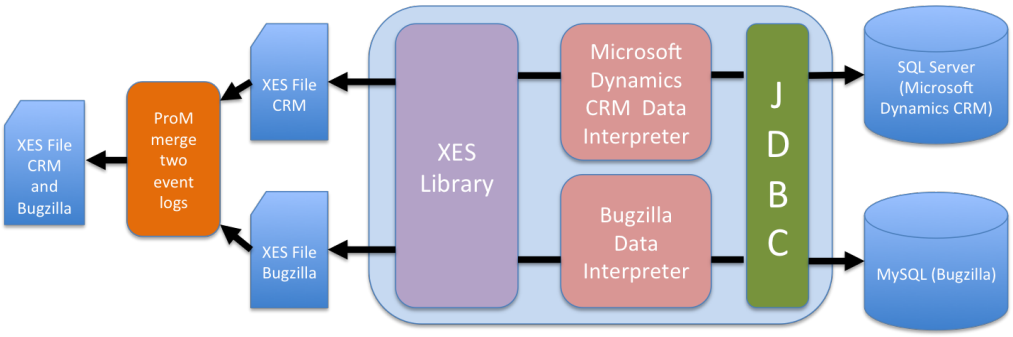


Diagram representing the different layers used to produce a single XES log file.

to execute can be measured. For the goal in which the project team has to deliver an understanding on how stakeholders are interacting in the Help Desk and Bugzilla processes we need to also include resources or people that are executing activities in the process. After a set of discussions, the team came with a list of desired attributes that are to be included for the data extraction task, this list is detailed in **section A.4** appendix.

4.2.1.1 Data Extraction

As discussed in the previous subsection, the data to be extracted from Microsoft Dynamics CRM and Bugzilla listed in **section A.4** appendix will be used so two XES log files via JDBC and a XES library are created. Once these XES files have been constructed, ProM 6's[27] *Merge two Event Logs using rule based algorithm* plug in is used to merge the two produced XES files. To merge both files CRM's *Case ID* and Bugzilla's *cf-rifcrm* data fields were selected to match cases using the aforementioned ProM 6 plug in. In **Figure 4.2** a diagram details the layering used to come up with the XES log file that will be used in our Process Mining project.

A selected case from our log file is detailed in **Table 4.1**, the structure of every case will allow the project team to scrutinize the log file during *DMAIC*'s *Analyze* phase. Note that not all cases involve the interaction

with Bugzilla process, in this example, a case has been selected which involved Bugzilla process activities so the data structure can be shown. Additionally, some information in **Table 4.1** has been anonymized per IT Company request.

TABLE 4.1: Selected Case from Log

Case ID	Activity	Resource	Complete Timestamp	bug_id	priority	product	seriousness	servicelevel	workgroup
CAS-12247-B6B0B4	Assign seriousness	Resource1	2/11/13 15:08			Product1	White Code	Basic	Helpdesk
CAS-12247-B6B0B4	Create SW Anomaly	Resource2	4/2/13 14:23			Product1	White Code	Basic	Helpdesk
CAS-12247-B6B0B4	Create SW Anomaly	Resource2	4/2/13 14:23			Product1	White Code	Basic	Helpdesk
CAS-12247-B6B0B4	IN_PROGRESS	Resource3	4/24/13 16:17	8138	Low				
CAS-12247-B6B0B4	RESOLVED	Resource3	4/29/13 16:04	8138	Low				
CAS-12247-B6B0B4	FIXED	Resource3	4/29/13 16:04	8138	Low				
CAS-12247-B6B0B4	CONFIRMED	Resource3	4/29/13 16:19	8138	Low				
CAS-12247-B6B0B4	RESOLVED	Resource3	4/29/13 18:13	8138	Low				
CAS-12247-B6B0B4	FIXED	Resource3	4/29/13 18:13	8138	Low				
CAS-12247-B6B0B4	Resolve SW anomaly	Resource4	5/2/13 8:05			Product1	White Code	Basic	Helpdesk
CAS-12247-B6B0B4	Resolve SW anomaly	Resource4	5/2/13 8:05			Product1	White Code	Basic	Helpdesk
CAS-12247-B6B0B4	Require upgrade	Resource5	5/2/13 8:51			Product1	White Code	Basic	Helpdesk
CAS-12247-B6B0B4	Resolve ticket	Resource2	5/15/13 10:53			Product1	White Code	Basic	Helpdesk
CAS-12247-B6B0B4	Closed	Resource6	5/30/13 10:53			Product1	White Code	Basic	Helpdesk
CAS-12247-B6B0B4	VERIFIED	Resource2	11/28/13 18:05	8138	Low				

Selected case showing Microsoft Dynamics CRM and Bugzilla systems activities, note that some data has been omitted and/or anonymized due to Company's request.

4.2.1.2 Log File Statistics Overview

After putting together the log file containing Help Desk's Microsoft Dynamics CRM and Bugzilla data, it was proceeded to explore the overall statistics and the process map produced by Disco™. This stage is of the Process Mining Project is working in synergy with *DMAIC*'s *Measure* phase, the goal of this stage is to communicate the findings detailed by Disco™ software to the team members so stakeholders can get familiarize with the information discovered by Disco's™ features.

The following set of figures aims at detailing the process definition, metric definitions, establish the process baseline, and evaluate the measurement system [20, p. 271]. These objectives set the basis for the project team to be able to establish a common language and understanding of the several metrics, Process Mining language, models and other relevant information that will be used as the project progresses.

In **Figure 4.3** a process map for both Help Desk and Bugzilla processes is displayed, here is relevant to mention that such process map discovered by Disco™ includes activities from two processes -Microsoft Dynamics CRM and Bugzilla- so it is clearly possible to successfully map a process with Disco™ using data from two processes. The interaction of two processes is relevant to understand their dynamics, in particular to understand whether the desired synchronization and safeguards documented by a de jure model are being executed as expected.

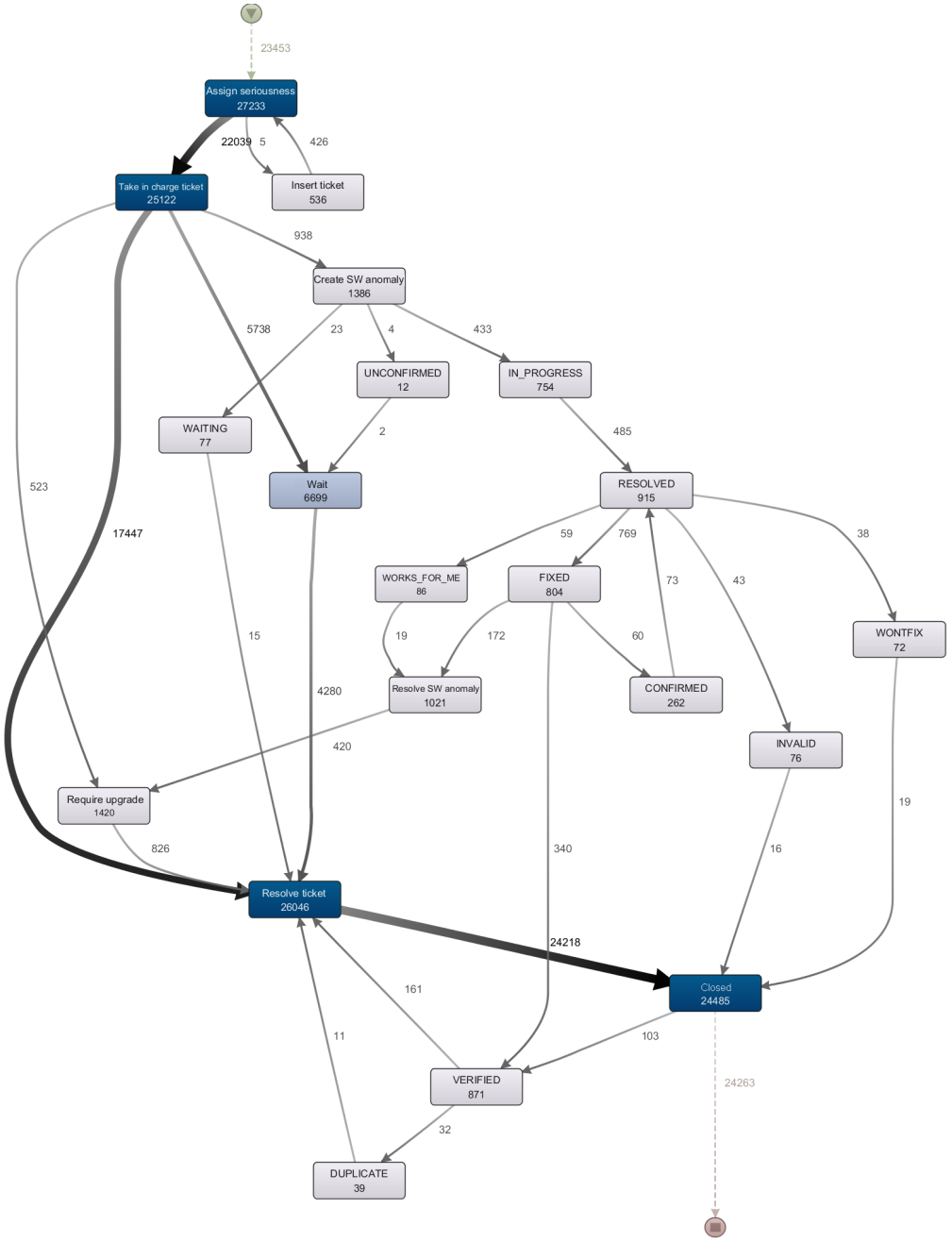
The events over time chart, from the *Overview* feature, shown in **Figure 4.4** as well as the events per case chart shown in **Figure 4.5**, are statistical data which clearly show the volume of activities being executed is increasing and that a typical Help Desk incident has about four activities executed for the typical process instance. The fact that the chart in **Figure 4.4** is showing that the number of activities being executed in a daily basis is increasing is an insight which in some extent helps to address item 6 detailed in **section A.3** appendix project's charter *Opportunity/Problem Statement* section: *There is a need to understand if the number of Help Desk incidents is increasing.* In the measurement of arrival of cases subsection of this chapter, another approach will be used to measure the increase of incidents over time.

In order to prepare for the task where a business process will be modeled based in interviews and a process map discovered by Disco™, the project team took note of the fact that the *Statistics* feature in Disco™ shows that the four most frequent activities are those same activities that are executed in the most frequent path. In **Figure 4.3** it is shown that the most frequent process instance pattern follows the sequence:

Assign seriousness→Take in charge ticket→Resolve ticket→Closed

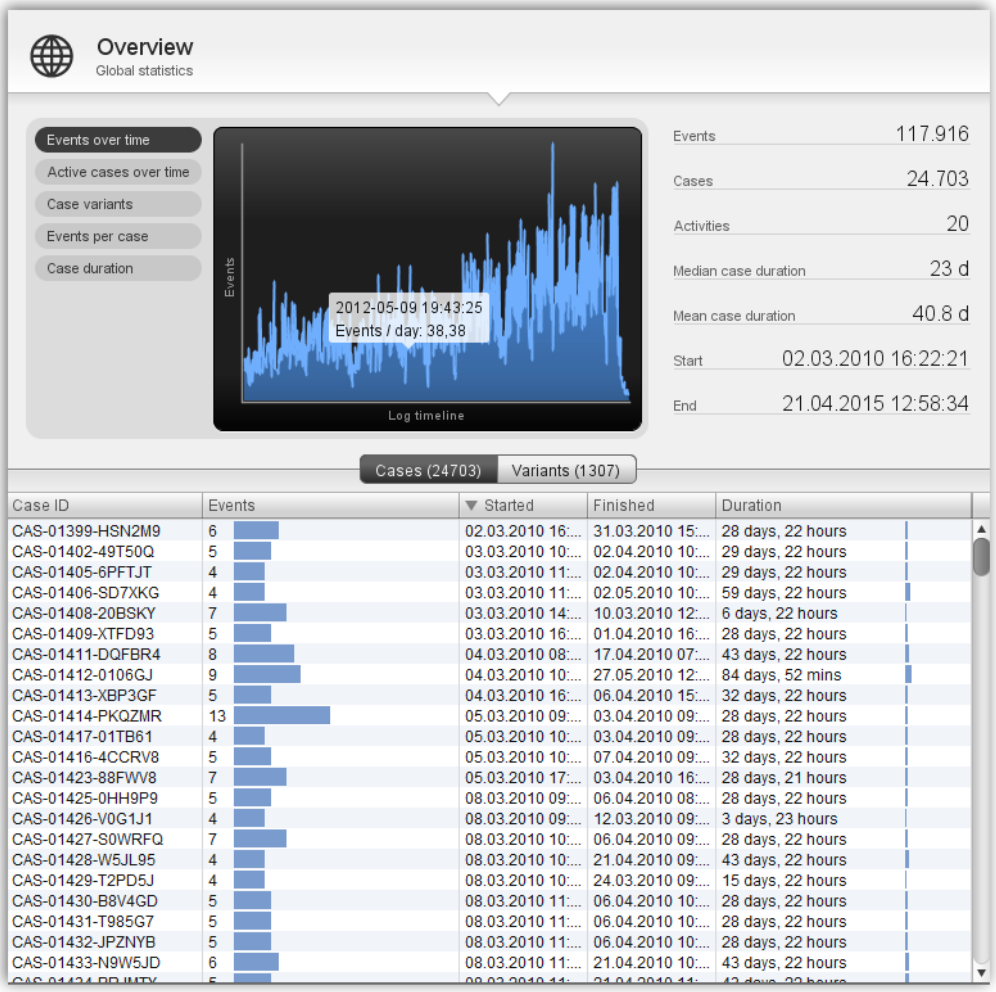
The discovery of this frequent path is going to be a fundamental piece of information that will aid in modeling a business process given the fact that activities

FIGURE 4.3: Microsoft Dynamics CRM Help Desk and Bugzilla Process Map



Process map discovered by Disco^{TM}, the map shows activities for Bugzilla and Help Desk systems. Disco^{TM} sliders are set to 100% and 1% for activities and paths, respectively.

FIGURE 4.4: Disco™ Overview Statistics: Events over time

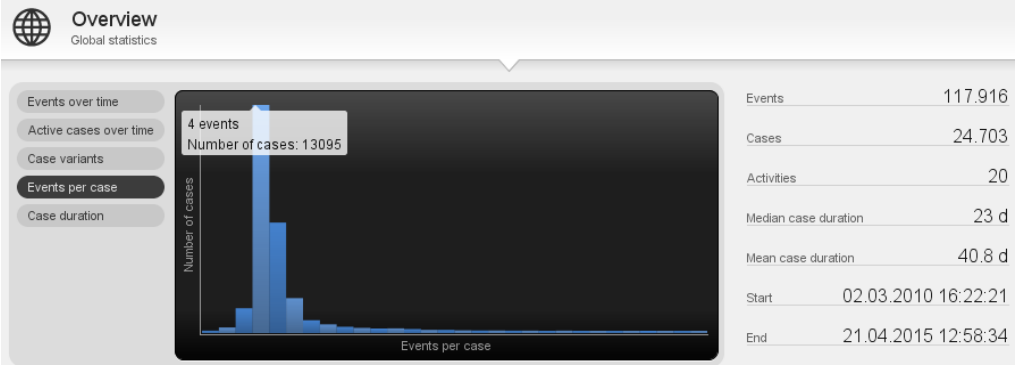


The events over time view from *Overview* statistics details the volume of activities being executed in a time series fashion.

involved in the most frequent path are also the most frequent activities detailed in Disco's™ *Activity event classes* view within the *Statistics* tab. **Figure 4.6** shows a detailed view in terms of activity events classes for Help Desk and Bugzilla processes.

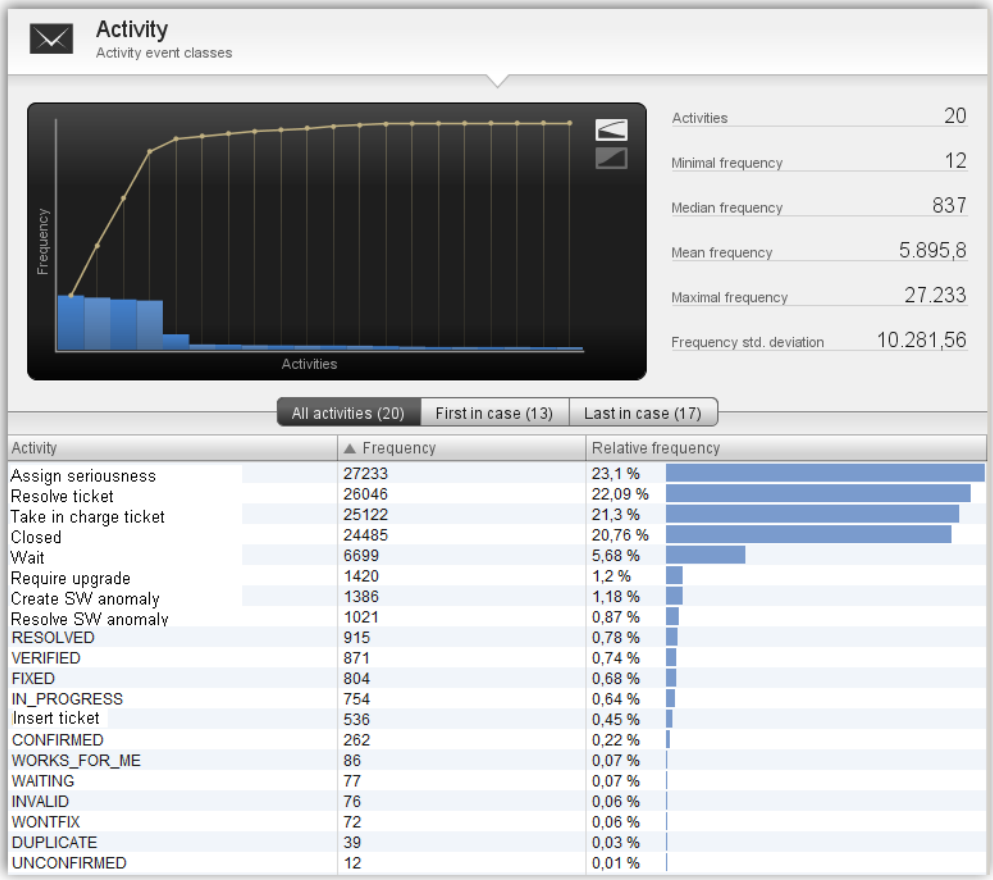
In terms of *Case variants*, Disco™ offers a view in which is possible to quickly visualize case variants in our Help Desk and Bugzilla processes, **Figure 4.7** illustrates this concept where we can realize that variants 1 to 13 account for approx. 84% of the variants in the discovered process. It is worth mentioning that none of these 13 variants involves activities from Bugzilla process so due to the low execution of Bugzilla activities, relative to Microsoft Dynamics CRM's

FIGURE 4.5: Disco^{TM} Overview Statistics: Events per case



The events per case view from *Overview* statistics details the number of events that each case has in the form of a histogram, a typical case has 4 events.

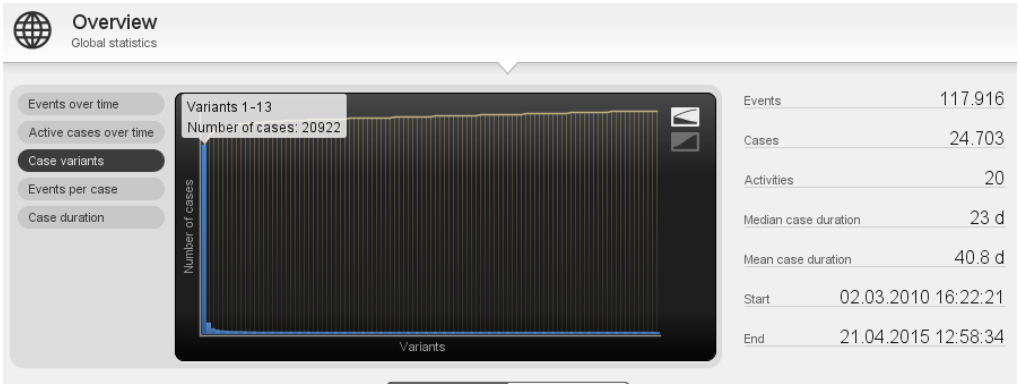
FIGURE 4.6: Disco^{TM} Activity Statistics: Activity event classes



Activity	Frequency	Relative frequency
Assign seriousness	27233	23,1 %
Resolve ticket	26046	22,09 %
Take in charge ticket	25122	21,3 %
Closed	24485	20,76 %
Wait	6699	5,68 %
Require upgrade	1420	1,2 %
Create SW anomaly	1386	1,18 %
Resolve SW anomaly	1021	0,87 %
RESOLVED	915	0,78 %
VERIFIED	871	0,74 %
FIXED	804	0,68 %
IN_PROGRESS	754	0,64 %
Insert ticket	536	0,45 %
CONFIRMED	262	0,22 %
WORKS_FOR_ME	86	0,07 %
WAITING	77	0,07 %
INVALID	76	0,06 %
WONTFIX	72	0,06 %
DUPLICATE	39	0,03 %
UNCONFIRMED	12	0,01 %

The *Activity event classes* view in Disco^{TM} shows that the four most executed activities are those that make the most frequent path as shown in Fig. 3.3.

FIGURE 4.7: Disco™ Overview Statistics: Case variants



The *Case variants* view in Disco™ shows that the 13 most common variants account for about 84% of the incidents in Help Desk and Bugzilla processes.

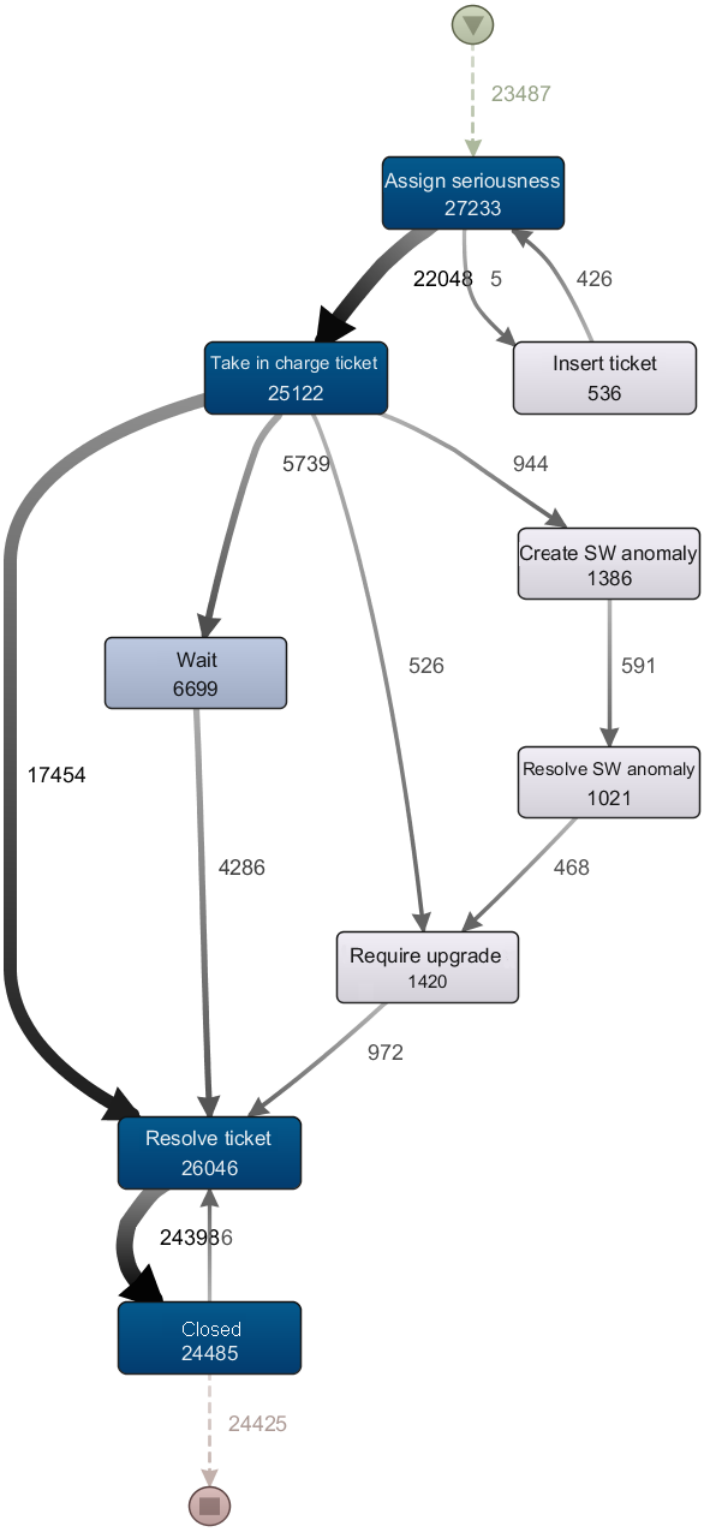
activities, it may be more practical to filter the log so only cases that involved Bugzilla system are analyzed.

4.2.2 Business Process Modeling based in Stakeholder Interviews and Process Mapping with Disco™

A set of interviews was conducted with Help Desk and Bugzilla processes stakeholders in order to map a handmade model that can be compared with a process map discovered by Disco™. During this task there ware about four iterations in which the handmade model and the map discovered by Disco™ were compared in order to refine a de jure model. The process map discovered by Disco™, which involves the Help Desk Process only, is detailed in **Figure 4.8**. This process map has been used during discussions and interviews to be able to build a de jure model for the Help Desk process.

After the hand made model has finalized, the team then proceeded to draft a model using BPMN 2.0 language. Here is worth mentioning that *DMAIC*'s *Measure* phase contemplates the documentation of a process [20, p. 272-273] in

FIGURE 4.8: Microsoft Dynamics CRM Help Desk Process Map



Process map discovered by Disco™ where only Microsoft Dynamics CRM activities are being shown. Disco™ sliders are set to 100% and 1% for activities and paths, respectively.

which the project team achieves this task by modeling the process detailed in the remaining of this subsection.

The Help Desk operation process is modeled in a main process and a sub-process, the main process is detailed in **Figure 4.9** top diagram. This top-level process shows how a simple ticket is first screened to decide whether there is enough information to assign a seriousness level to the incident, if there is not enough information to determine the incident's seriousness the ticket is then logged as *Insert ticket*. Once there is enough information to determine the seriousness of the incident the Help Desk team member assigns a seriousness level to the incident by executing activity *Assign seriousness*. There are three seriousness levels:

- Level 1 red code
- Level 2 yellow code
- Level 3 white code

The seriousness levels detailed above have been discussed in the SLA table section.

Once an incident has been assigned a seriousness level, a Help Desk resource can resolve the customer's incident by further investigating the incident. If the ticket requires further investigation it will be resolved within a ticket investigation sub-process, this sub-process is detailed in the following paragraph.

The ticket investigation sub-process is detailed in **Figure 4.9** bottom diagram, note that once the ticket has been assigned a seriousness level, it is then processed to be reviewed. The ticket investigation sub-process starts when the ticket is assigned an owner by executing activity *Take in charge ticket*, once this happens, a resource assigned to the case starts an investigation. Here, the Help Desk resource may decide, or not, that further information is needed to continue the investigation, if that is the case, the resource executed the activity *Wait* indicating that the state of the case moves to a waiting state. Here is important to point out that once the ticket is set to a waiting state, the SLA counter pauses,

so while the customer is informed that more information is needed, the time the customer takes to respond does not count toward SLA.

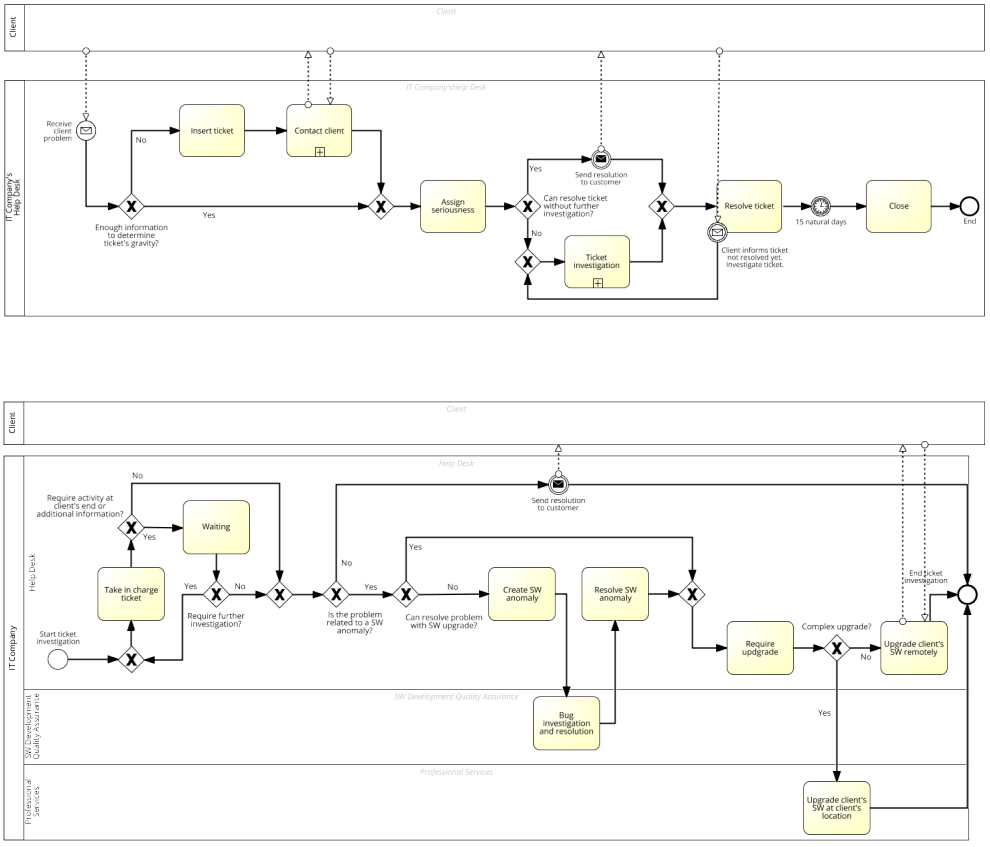
Once it has been determined that all the required information is available, and no further investigation is needed, the case then can be categorized as one that requires the customer to upgrade its software, send a resolution to the customer, or that the software application being serviced has a potential anomaly which requires the software development team intervention. Here, because software anomalies require time to be fixed (in case the SW development team decides to address the anomaly) the Help Desk resource logs a case in a separate system called Bugzilla. At this point the Help Desk team will try to address customers needs by resolving the issue with a work-around while the SW development team addresses the potential bug.

In case the customer is asked to upgrade the SW, it is possible to do it remotely or at the customer's location. Once the case has gone through the sub-process detailed above, the case then exits the ticket investigation sub-process and re-enters the main process. Here once the customer has been notified about the possible resolution the incident moves to *Resolved* state,at this point, the the information system starts a countdown of 10 business days. If the customer reports within these ten business days that the provided resolution did not address the initially reported incident, the case then returns to the ticket investigation sub-process. If the customer does not contact the Help Desk operation within these ten business days, the case is automatically closed by executing the activity *Closed*.

4.2.3 Measuring Arrival of Cases for the Help Desk Operation

In our project charter, the team set as one of the goals to understand *how the different software releases are influencing the number of Help Desk incidents related to potential SW anomalies issues.* For this, the team determined that there

FIGURE 4.9: Company's Help Desk main process and ticket investigation sub-process.



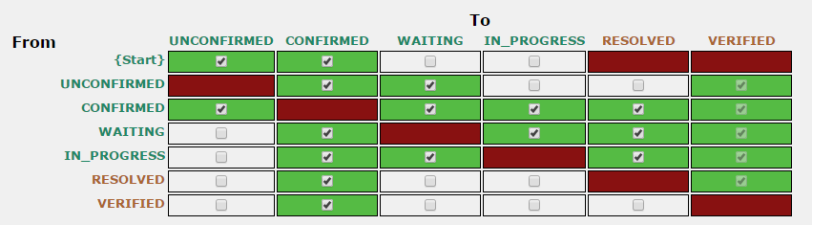
The main process and ticket investigation sub-process, top and bottom diagram, respectively.

has to be a measurement of the amount of incidents being logged in Microsoft Dynamics CRM process as a function of time. This task can be accomplished with ProM 6 *Xdotted Chart plug in*, this plug in shows the arrival of cases or tickets and its subsequent activities, this is, the chart shows all activities being executed in the log where the y axis denotes the different cases logged in the system and the x axis describes when an activity was executed based in activities' time stamp. Appendix in **section A.5** details the arrival of cases, here it can be seen that some batching is happening in the process, this is indicated by the same activities executed in a short period of time and for different cases. This batching is annotated with the pink ovals. Also, note in this chart how the arrival of cases is increasing over time.

The *XDotted* chart accomplishes one of the project charter tasks by measuring the volume of incidents being logged in the CRM process for which the following chapter we will discuss its analysis to understand how software releases are impacting the volume of incidents that involve potential SW anomalies.

4.2.4 Mapping the Bugzilla Process

FIGURE 4.10: Bugzilla de Jure Model as a Table

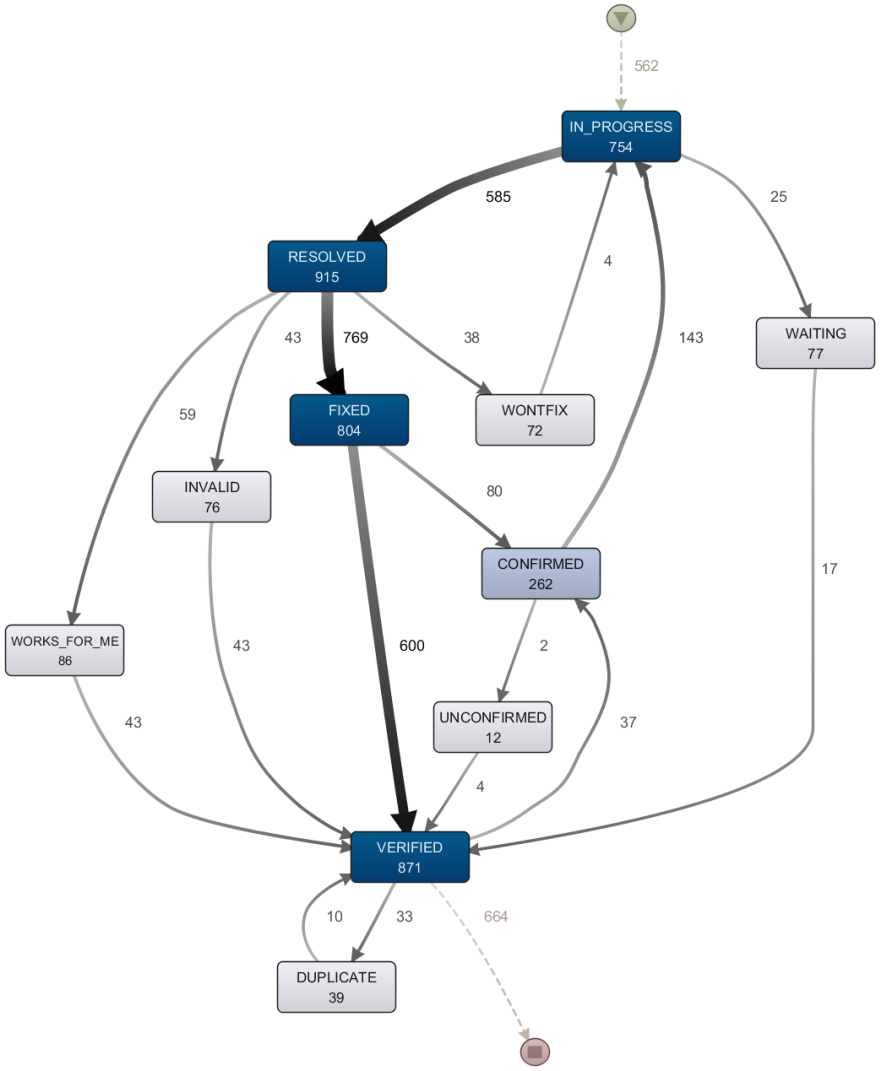


Since Bugzilla is a stand alone system to manage software anomalies, the team considered as part of the tasks to map a de facto process model with Disco™ Process Mining software to later on compare it to a de jure model, the purpose of this task is to be able to measure in what degree the de jure model is being executed accordingly. The de jure model for Bugzilla is depicted **Figure 4.10**, the process that has been mapped with Disco™ is shown in **Figure 4.11**.

This table details the transitions that are allowed between activities, the red cells indicate transitions are not allowed, green cells indicate transitions are allowed, and white cells indicate transitions are allowed but have been disabled.

Even though a de jure model for Bugzilla process has been represented with a table, it is possible to deduct whether the de facto model synchronizes with the de jure model, for instance, is possible to have the *IN_PROGRESS* activity followed by *RESOLVED* activity which can be succeeded by *FIXED* activity, therefore, the de facto model should show the same behavior as the transitions shown in **Figure 4.11**. A more detailed model for Bugzilla is described in Bugzilla's guide documentation [28].

FIGURE 4.11: Bugzilla Process Map



Snapshot from Disco^{TM} software application showing only activities from Bugzilla application, Microsoft Dynamics CRM process activities are not shown.

4.2.5 Mapping the Organizational Perspective

An understanding on how entities, in terms of human resources, interact within processes is a goal that has been outlined in the project charter. To answer this question there is a need to map the organizational perspective in terms of handover of work, the goal is to visualize how Help Desk cases are being handled between resources working the Help Desk process and from resources in the Help Desk team to resources in the SW Quality Assurance team. The

organizational map discovered by Disco™ is detailed in **section A.7**, this map has been constructed, in part, by following instructions in Fluxicon™'s blog entry[29, p. 38].

As it can be deducted from the map, *HD Res1* seems to be a central part in the process, after consultations with the Help Desk manager it was confirmed that this resource performs roles in managerial capacity so cases are being handed over by this resource to other resources in the Help Desk team, i.e. *HD Res3, HD Res6, HD Res9*. Similarly, cases assigned to the previously mentioned resources are consequently handed over to the SW Quality Assurance team where for instance *SW QA Res4* is the resources handling most of the incidents that at some point during the process were identified as potential SW anomalies.

It is worth mentioning that the organizational map constructed with Disco™ has been filtered in such a way that only Help Desk incidents involving potential SW anomalies are being considered, incidents that are not considered as potential SW anomalies are being discarded from the log used to construct the organizational map. At this stage of in *DMAIC*'s *Measure* phase, the organizational map brings a good insight on how cases are being handed between resources so this information can be used to propose improvements if needed.

CHAPTER 5

Extending *L* Life-cycle Model* Stages 2 and 3: Create Control-Flow Model, Connect Event Log, and Create Integrated Process Model

In the previous chapter we discussed how a Process Mining Project via the *L* model* in its *extract* stage can be extended with *DMAIC*'s *Measure* phase activities such as defining a more detailed model and establishing a measurement method by extracting the relevant data and creating an XES file so a log file can be properly read by Disco and ProM. In **section A.4** appendix we also defined our fields or attributes for the log file that will serve the purpose of paving the way for the *Analyze* phase.

In terms of extending the *L* model* with a Quality Management System we discussed how the *extract* phase of the *L* model* can be enriched in selecting and designing the appropriate KPIs in accordance with the ISO 9001:2008 QMS.

With these accomplishments in hand, in this chapter we will discuss how the *L* model* can be extended with *DMAIC*'s *Analyze* phase and will discuss how the *L* model* can be extended with ISO 9001:2008 QMS audit requirements and how these findings can provide feedback to management in order to address improvement areas in the process or processes in question.

In terms of EFQM Excellence Model™ and Baldrige Criteria™ frameworks, we will discuss how the *L* model* can use these frameworks assessments in order to diagnose the needs of the organization so appropriate questions can be formulated which can be translated in having more clarity in what are the goals for stages 2 and 3 of the *L* life-cycle model*.

5.1 Organizations Resources and Quality Management System and Business Improvement Frameworks

Stage of 3 the *L* model*, *Create Integrated Process Model*, consists in adding additional perspectives to the control flow model acquired during stage 2. Some of these additional perspectives are the *organizational perspective*, *case perspective*, or *time perspective*. For this subsection of the chapter, we will discuss how the organizational perspective in stage 3 of the model can be expanded to include criteria from the EFQM Excellence Model™ and the Baldrige Criteria™.

The EFQM Excellence Model™ details in its criteria *4. Partnerships & Resources* section *4a* that *"Technology is managed to support the delivery of strategy"* in the context of *"identifying and evaluating alternative and emerging technologies in the light of the impact on organizational performance and capabilities"*[13, p. 14]. As well, the same section mentions that excellent organizations *"involve relevant stakeholders in the development and deployment of new technologies to maximize the benefits generated"*[13, p. 14]. The concepts detailed in these criterion can be interpreted in the context where a new technology exploiting Process Mining techniques can be adopted in order to better involve the stakeholders in key processes.

The EFQM Excellence Model™ also details in its *3. People* criteria that in excellent organizations *"people plans the support of the organization's strategy"* by aligning *"people plans with their strategy, the organizational structure, new technologies and key processes"*[13, p.]. The criteria also mentions that excellent organizations *"align, involve and empower people"* by *"recognizing that innovation can apply to processes and organizational structures."*[13, p. 13] In the light

of these statements and as mentioned in the previous paragraph, organizations that use technologies that adopt Process Mining techniques through the use of the *L* model* can achieve in meeting the criterion discussed earlier. Which consequently expands the *L* model* in its stage 2.

Concerning the Baldrige Criteria^{TM}, the *5 Workforce* criteria in its *Workforce Environment* section details in paragraph *a4* that in terms of *"Workforce Change Management"*[17, p. 21] organizations need to question whether the organization is preparing the workforce for change in *"capability and capacity"* needs in terms of *organizational structure and work systems as needed*. This criteria can be used to expand the *L* model* in its stage 3, so the aforementioned principles can be considered during a Process Mining project.

Additionally, the Baldrige Criteria^{TM} states in its *Operations* section in *Process management* that organizations need to question whether in terms of process implementation *"work processes meet key process requirements in a daily basis"*[17, p. 23]. Similarly, the criteria states that organizations need to question whether *"work processes are being improved so products and performance are being developed"*.

The above mentioned criterion from the Baldrige Criteria^{TM} can also be integrated into the stage 3 of the *L* model* so the creation of an integrated process model considers such principles.

5.2 Extending Process Mining *L* Life-cycle Model* Stages 2 and 3 with Six Sigma's *Analyze* Phase

Stages 2 and 3 of the *L* model* for Process Mining projects aim at determining the de facto control-flow model of the process being analyzed in case a process model has not being determined, if there is a de jure model available it is possible to perform conformance checking diagnostics[5, p. 285-286]. Given that in our Process Mining project we have modeled a process via interviews and with the help of Disco^{TM} process mining software, we will analyze the log to find sources

of variation, trends and changes in incidents arrival to the process using Process Mining techniques in conjunction with *DMAIC*'s *Analyze* phase.

5.2.1 Help Desk Process Conformance Checking Diagnostics Analysis

Because *DMAIC*'s *Analyze* phase aims at –among otter things– finding sources of variation, in this section we will perform conformance checking diagnostics over periods of six months in order to understand in what measure the overall process trace fitness is performing. For this task an analysis was performed under the following methodology:

(**a**) A simplified version of the Help Desk process has been modeled in BPMN 2.0 language so it is easier to convert it to a Petri net, this simplified version is detailed in **Figure 5.1**.

(**b**) The Help Desk process was modeled in the form of a Petri net that can reproduce the same behavior per the simplified Help Desk process detailed in Figure 5.1, so such Petri net can be used to perform conformance checking diagnostics using ProM 6 plug ins. A Petri net modeling the Help Desk is detailed in **Figure 5.2**.

(**c**) Once a Petri net representing the simplified Help Desk process has been constructed, we proceed to filter our log so we can aggregate incidents in periods of time. Here we use Disco^{TM} software to filter the log by time frame where we select cases that started in a given time frame window. This task achieves the purpose of aggregating cases based in time frames based in the first six months or last six months of a given year, that is, aggregating cases based in semesters. **Figure 5.3** details an example of one of the *time frame* filters used, note that the all activities for cases that started in a given time frame are being included. The sub grouping of the original log resulted in 11 logs for a time period ranging from Jan 2010 to May 2015.

(**d**) Once we have aggregated log cases based in *time frames* we proceed to run conformance checking analyses, here we used two plug ins:

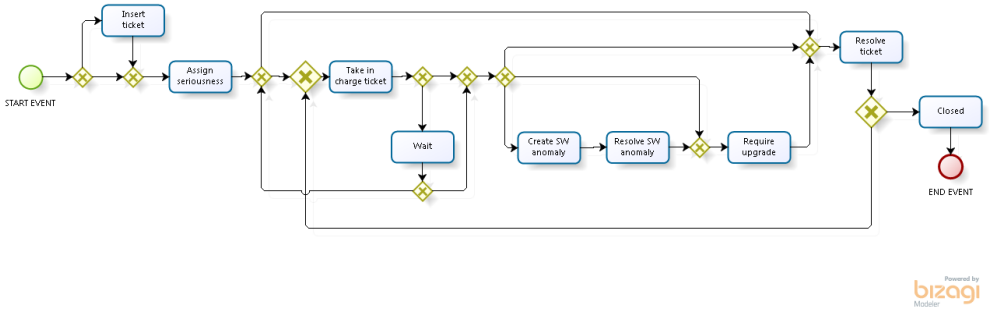
- ProM 6 *Replay a Log on Petri Net for Conformance Analysis* [30], and
- ProM 6 *Conformance Checking of a Petri Net with Data (results on BPMN)* [31].

The first plug in produces a Petri net with conformance checking diagnostics based in the de jure model detailed in Figure 5.2 where the *Trace Fitness* value from *Global Statistics* is used to note the overall process' fitness. **Figure 5.4** shows the Petri net output and conformance checking diagnostics.

The second plug in produces a BPMN 2.0 diagram also based in Figure 5.2 Petri net, running this plug in results in conformance checking diagnostics for every activity involved in the process, it also details the sequence of activities for those traces that had activities missing and/or activities not predicted by the de jure model. **Figure 5.5** illustrates the output for this plug in, in **sub-figure 5.5 a)** the coloring for each activity denotes how well synchronized each activity is given the de jure model in Figure 5.2, green activities are in better sync with the de jure model than activities colored in, yellow, orange, or red, respectively. **sub-figure 5.5 b)** details the synchronization diagnostics for traces having missing or unexpected behaviour given the de jure model, activities in purple are missing in the log, activities in yellow are activities observed in the log but not predicted by the model.

(**e**) Once we have acquired conformance checking diagnostics values we proceed to compile this information by structuring our data, appendix in **section A.6** details how our data is structured. The data structure aims at producing time series analyses charts and summary statistics charts such as box plots. Because *DMAIC*'s *Analyze* phase works in finding sources of variation [20, p. 427], analyzing our data structure as a function of time and fitness measure results in discovering where the variation is

FIGURE 5.1: Simplified Help Desk Process



This Help Desk process in BPMN 2.0 language is the simplified version of our Help Desk process detailed in Figure 4.9, a simplified process is needed to build a Petri net which will serve as a de jure model to perform conformance checking diagnostics with ProM 6 plug ins.

originating. Note that the *Fitness Measure* calculations were acquired the following way:

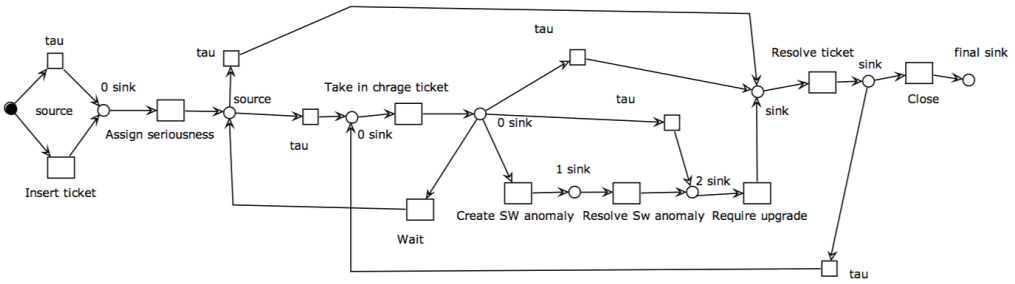
- For the *Overall process* fitness measure, this was gathered with *Replay a Log on Petri Net for Conformance Analysis* plug in result from its *Global Statistics Trace Fitness* as detailed in Figure 5.4.
- For the rest of the fitness measures –*Insert ticket, Assign severity, Take in charge ticket, Wait*, etc– we calculated the fitness measure by dividing the addition of skipped activities (***SAs***) and unexpectedly executed activities (***UEAs***) over the addition of executed as expected activities (***EEAs***), skipped activities (***SAs***) and unexpectedly executed activities (***UEAs***) where this was subtracted to an unit:

$$fitness\_measure = 1 - \frac{SAs + UEAs}{EEAs + SAs + UEAs}$$

The notion of skipped activities, unexpectedly executed activities, and executed as expected activities is detailed in *Conformance Checking of a Petri Net with Data (results on BPMN)* plug in output as illustrated in Figure 5.5, once clicking on one of the BPMN diagram activities, conformance checking diagnostics are shown in the pane.

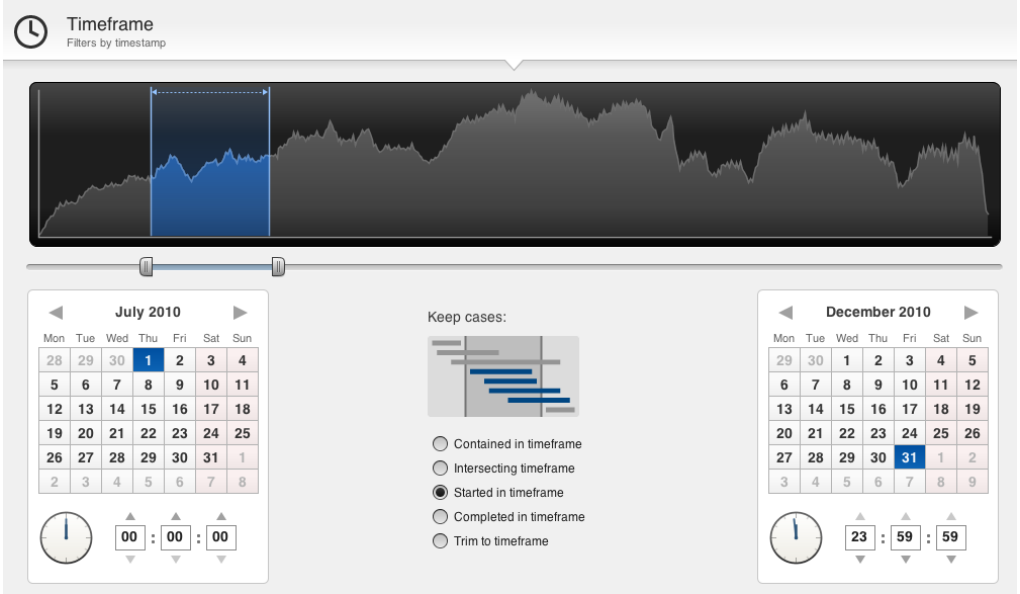
Given the above methodology and after getting the resulting data set to analyze causes of variation, we can now proceed to chart plots that help the project team

FIGURE 5.2: Help Desk Process as a Petri Net



This Petri net aims at modeling the Help Desk process detailed in Figure 5.1, it will be used to perform conformance checking diagnostics with ProM 6 plug ins.

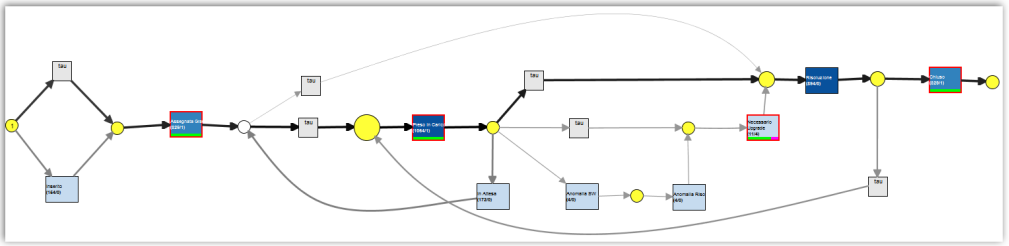
FIGURE 5.3: Time Frame Filter with Disco



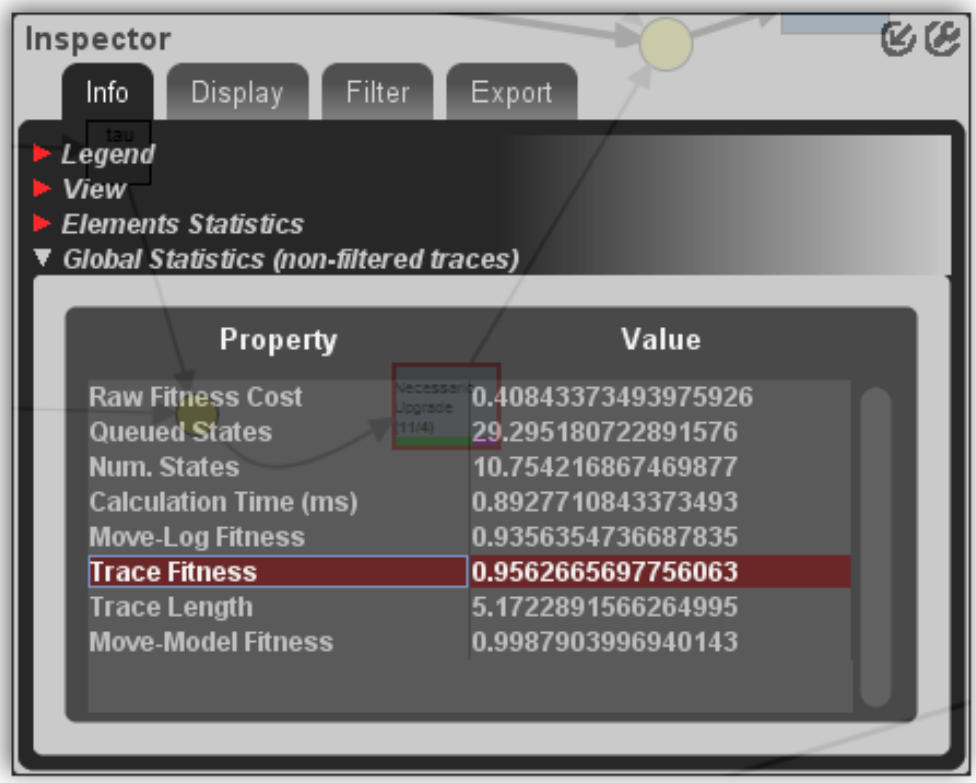
The time frame filter with Disco software serves the purpose of selecting cases that start within a given time frame so later on cases can be analyzed for conformance checking diagnostics to be able to analyze the process conformance to the de jure model in a time series approach.

FIGURE 5.4: Replay a Log on Petri Net for Conformance Analysis plug in output.

a)



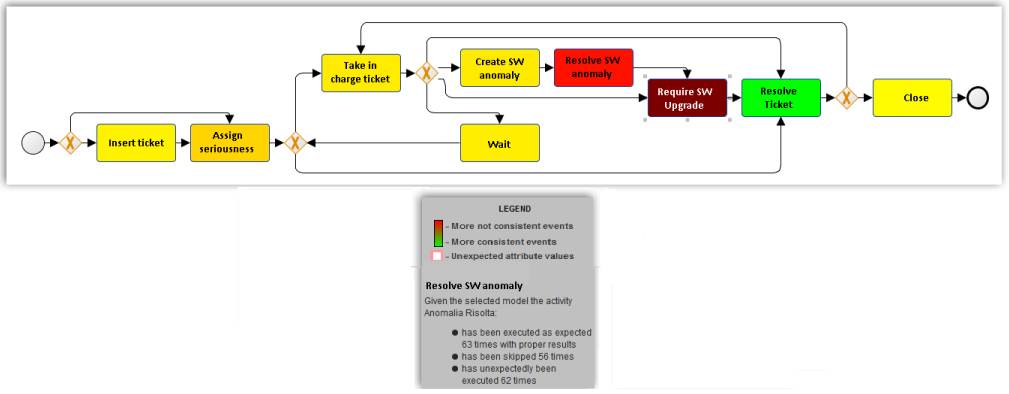
b)



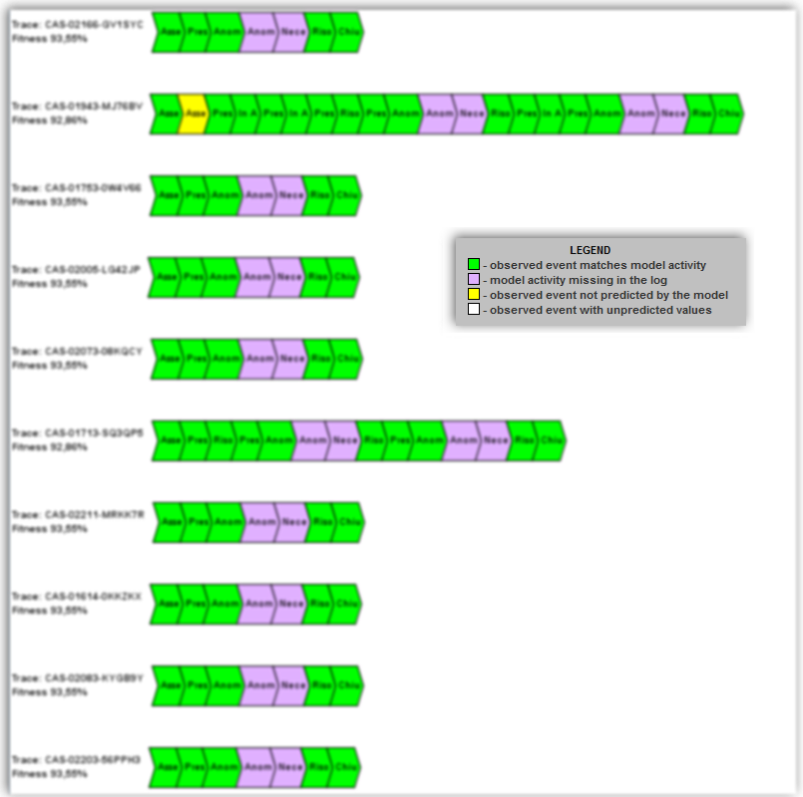
Sub-figure a) shows the plug in output, the Petri net used to run this ProM 6 plug in is the one detailed in Figure 5.2. In sub-figure b) the *Trace Fitness* value is highlighted in red.

FIGURE 5.5: Conformance Checking of a Petri Net with Data (results on BPMN) plug in output.

a)

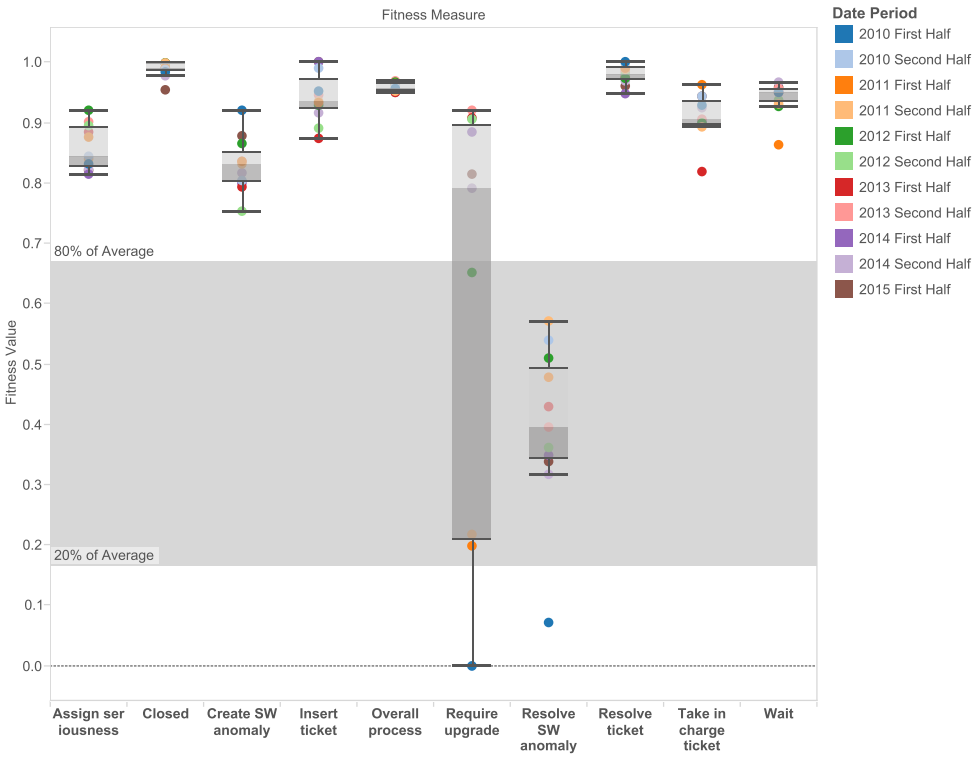


b)



Sub-figure a) shows PROMPT plug in output in the form of BPMN diagram, the Petri net used to run such ProM 6 plug in is the same as the one detailed in Figure 5.2. The definitions for the legend in this sub-figure are: executed as expected activities (***EEAs***), skipped activities (***SAs***) and unexpectedly executed activities (***UEAs***). In sub-figure b) it is detailed some traces that had missing or unexpected behaviour per the de jure model. Sub-figure b) has been blurred out in order to anonymize Company's proprietary data.

FIGURE 5.6: Help Desk Process Conformance Checking Fitness Box Plots with Outliers

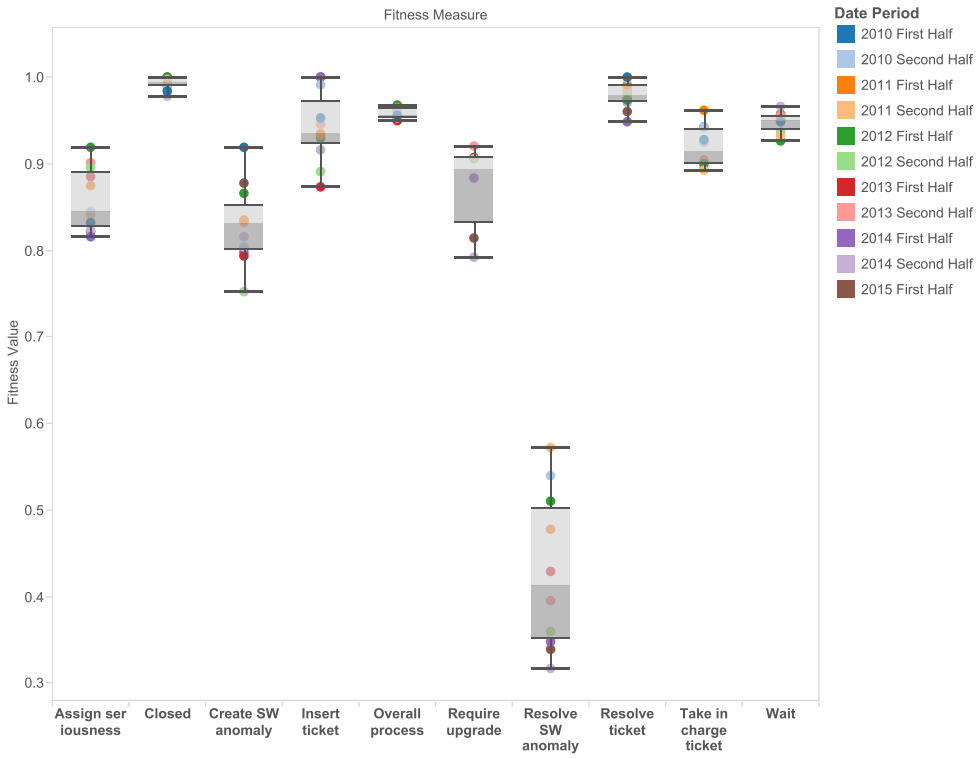


The box plots in this figure which depict summary statistics from fitness value and fitness measure for each date period clearly show that *Resolve SW anomaly* and *Require upgrade* activities need to be scrutinized due to systematically low fitness value and high fitness variation, respectively.

identifying sources of variation, for this, Tableau business intelligence software [32] was used in order to build box plots as a function of fitness value, fitness measure and time period. **Figure 5.6** details the resulting box plots charts created with Tableau software, here it can be noted that the chart illustrates that activity *Resolve SW anomaly* has been, in a systematic way, having lower fitness measure values when compared to the rest of the activities.

Since the box plots in Figure 5.6 have outliers, it is necessary to remove these outliers to get box plots without noise. For this, **Figure 5.7** illustrates that after outliers removal, *Resolve SW anomaly* activity remains an activity with consistently lower fitness measure values. By digesting the information detailed in Figure 5.7, it can be concluded that the process in which incidents are classified

FIGURE 5.7: Help Desk Process Conformance Checking Fitness Box Plots without Outliers



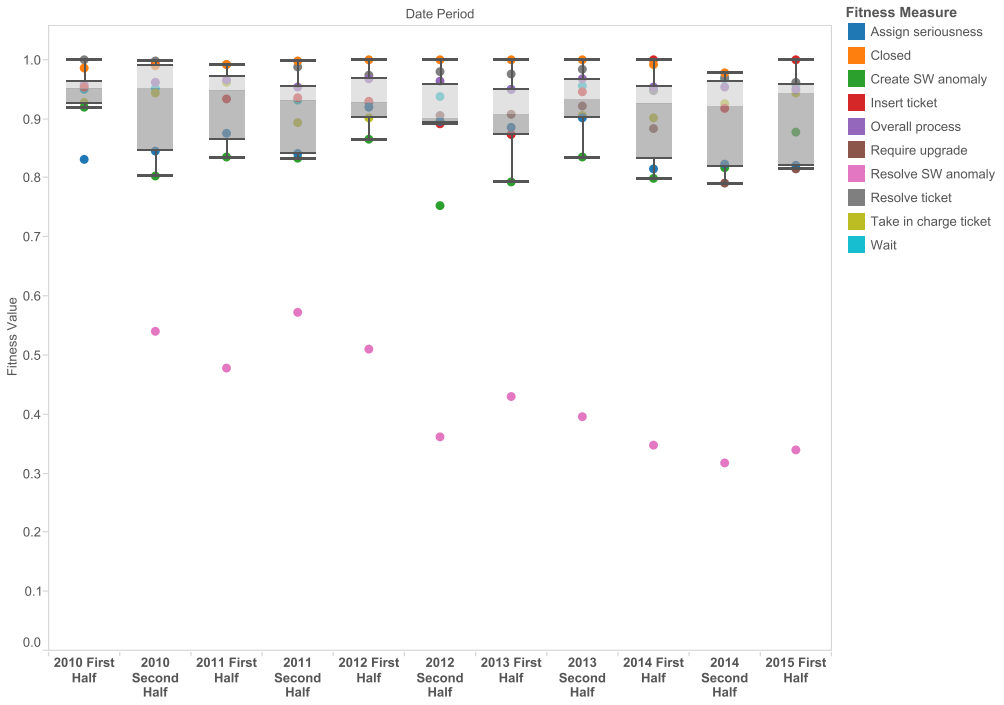
The box plots in this figure are now detailed without outliers, it can be noted that *Resolve SW anomaly* activity remains with a lower conformance fitness value for all semesters in which it was measured.

as those where there is a potential software anomalies have, systematically, lower conformance checking fitness values.

This finding has prompted the task to re-evaluate the actual business process for which a proposal for a change will be made, so the business process can benefit of such changes in order to improve the conformance checking fitness value for *Resolve SW anomaly* activity. These proposals will be discussed in the succeeding chapter.

An additional perspective is to analyze fitness measurement values as a function of time and fitness value, for this task an additional box plot chart was created, in contrast to the box plot detailed in **Figures 5.6** and **5.7**, the box plots detailed in **Figure 5.8** show that activities *Require upgrade* and *Resolve SW anomaly* have considerable lower fitness values. *Require upgrade* activity has lower fitness

FIGURE 5.8: Help Desk Process Conformance Checking Fitness Box Plots as Time Series



The box plots in this figure detail how several activities perform over time in terms of fitness value for the activities in the Help Desk process.

values between the first half of 2010 and the first half of 2012. This chart hints the fact that there has been a process change in which activity *Require upgrade* was not part of the process for the year of 2010, similarly, this same activity started to be used in a more normal fashion during the following 18 months to later on be used fully normally for the rest of the periods detailed in the chart.

Here it is worth stressing out that by measuring conformance checking in the form of fitness for every activity helps in identify process changes in terms of whether activities are being added, removed or changed in regards to a current or obsolete de jure model. Therefore, by measuring the variation of conformance checking fitness values is possible to identify activities in a process that may need further supervision, so activities with lower fitness values can be easily identified to be brought to desirable conformance checking fitness value.

5.2.2 Help Desk Incidents Time Perspective and SW Release Versions Impact Analysis

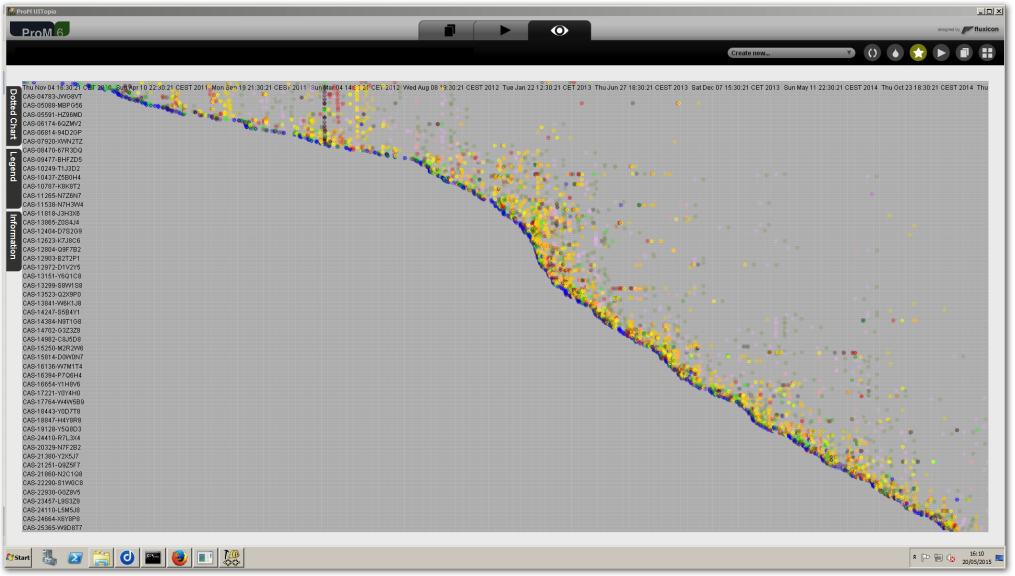
During some of the discussions with Help Desk staff a question regarding arrival of cases and workload was raised. For this, it is possible to analyze this phenomenon in two ways, the first one is to use ProM6 XDotted chart and the second one is using Disco's *statistics-events-over-time* view. Disco's solution approaches the workload question taking in consideration the events being executed given the time stamp for every event. As we discussed in in **section A.5** appendix, ProM 6 approach shows the arrival of cases or tickets and its subsequent activities, this is, the chart shows all activities being executed in the log where the y axis denotes the different cases and the x axis describes when an event happened. In section A.5 appendix is also shown some batching happening in the process, this is indicated by the phenomenon in which the same activities are being executed in a short period of time for different cases or ticket incidents. This batching is annotated with the pink ovals. Also, note in this chart how the arrival of cases is increasing, indicating an increase in workload. By understanding whether batching is happening or not in a process, process owners can realize whether certain resources may not be executing processes as desired.

The measurement of cases arrival in terms of time perspective needs now an analysis so item 5 in our problem statement project charter can be addressed: *There is a need to understand how the different software releases are influencing the number of Help Desk incidents related to potential SW anomalies issues.* To address this task we first need to consider Help Desk incidents that involve potential SW anomalies and that in consequence required the creation of a Bugzilla process instance, so to achieve this we filtered Help Desk incidents which had activity *Create SW anomaly* executed at least once, then the resulting log was plotted using *XDotted* chart plug in, the resulting chart is shown in **Figure 5.9 a)**.

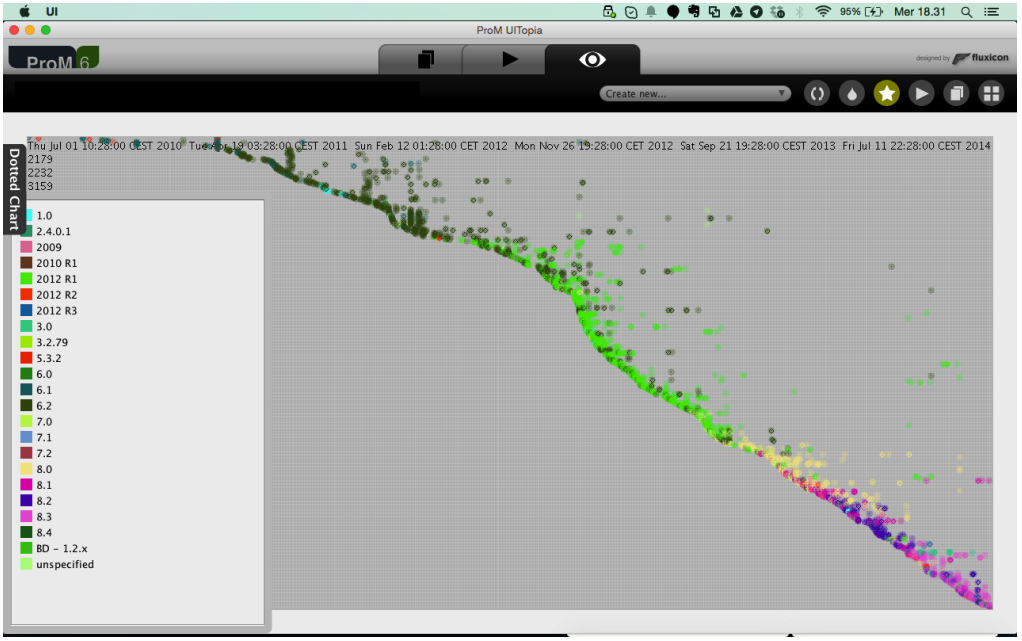
Given the fact that it is now understood there was a considerable increase in the number of incidents involving SW anomalies –note the increase of process

FIGURE 5.9: ProM 6 XDotted Charts with Cases Involving Bugzilla Interaction

a)



b)



Sub-figure a) shows the arrival of cases for company's main product that involve the interaction with Bugzilla system, this is, incidents that at some point involved the execution of activity *Create SW anomaly*. In sub-figure b), it can be noted the version of the main product that caused the increased activity of incidents involving potential SW anomalies.

instances in the center of the chart– we now need to find out what SW release caused such increase. In order to answer this question we use XDotted chart *Data type color* option and set it to show coloring as a function of the SW release version, **Figure 5.9 b)** shows that software release *2012.R1 6.10X* was installed at the customers' end some time between the actual release and when the increase of tickets involving interaction with Bugzilla started to happen around the beginning of February 2013. In this context, at some point in the last week of April 2013 the arrival of cases normalizes again.

This analysis, where time perspective is used to analyze how the amount of process instances involving *Create SW anomaly* activity increases, is relevant for management to asses whether performance indicators may be on target.

5.2.3 Service Level Agreement Metrics Analysis

As it has been noted earlier, The IT Company has service level agreements with its customers which oblige it to resolve incidents reported by clients in certain periods of time depending on the severity of the incident and service level purchased by the customer. In this context, we will discuss to what extend incidents are being resolved. For this it is important to note that while there is an intended work-flow that must be followed, such work-flow implemented in Microsoft Dynamics CRM lacks the necessary safe guards to confine the decision makers to adhere to the intended model. Therefore, we will analyze the duration of cases under the following conditions:

- Cases that do not start with *Insert ticket* or *Assign seriousness* will not be considered for this analysis.
- Cases that do not have *Insert ticket*, *Assign seriousness*, *Take in charge ticket* and *Resolve ticket* activities are also being excluded from the analysis.

The above conditions serve the purpose of filtering cases that do not follow a work flow that can be used to calculate the duration between 'Assign severity' to

'Presso in Carico' and from *Assign seriousness* to *Resolve ticket*. These durations are in essence the duration between these activities so it can be understood in what extent such durations are complying with SLA. Because most of the cases have *Assign seriousness* as the first event, we will also not consider the duration between *Insert ticket* to *Assign seriousness* due to the fact that from the 24 612 cases, only 503 cases have *Insert ticket* activity as a first activity.

Another aspect to take in consideration is that in some cases some activities are executed more than once, for instance, case *CAS-13521-F3T7K1* has the following set of activities which are sorted as a function of time stamp in ascending order:

TABLE 5.1: Repeated Activities in Selected Case

CAS-13521-F3T7K1	Assign seriousness	2013/05/03 09:48:23.000
CAS-13521-F3T7K1	Assign seriousness	2013/05/20 08:22:15.000
CAS-13521-F3T7K1	Assign seriousness	2013/05/20 08:22:17.000
CAS-13521-F3T7K1	Resolve ticket	2013/05/21 06:56:12.000
CAS-13521-F3T7K1	Resolve ticket	2013/05/21 06:56:13.000

Table showing repeated activities for the same case, when the same activity is executed more than once, the first instance of that activity is the one being considered for the SLA measurement.

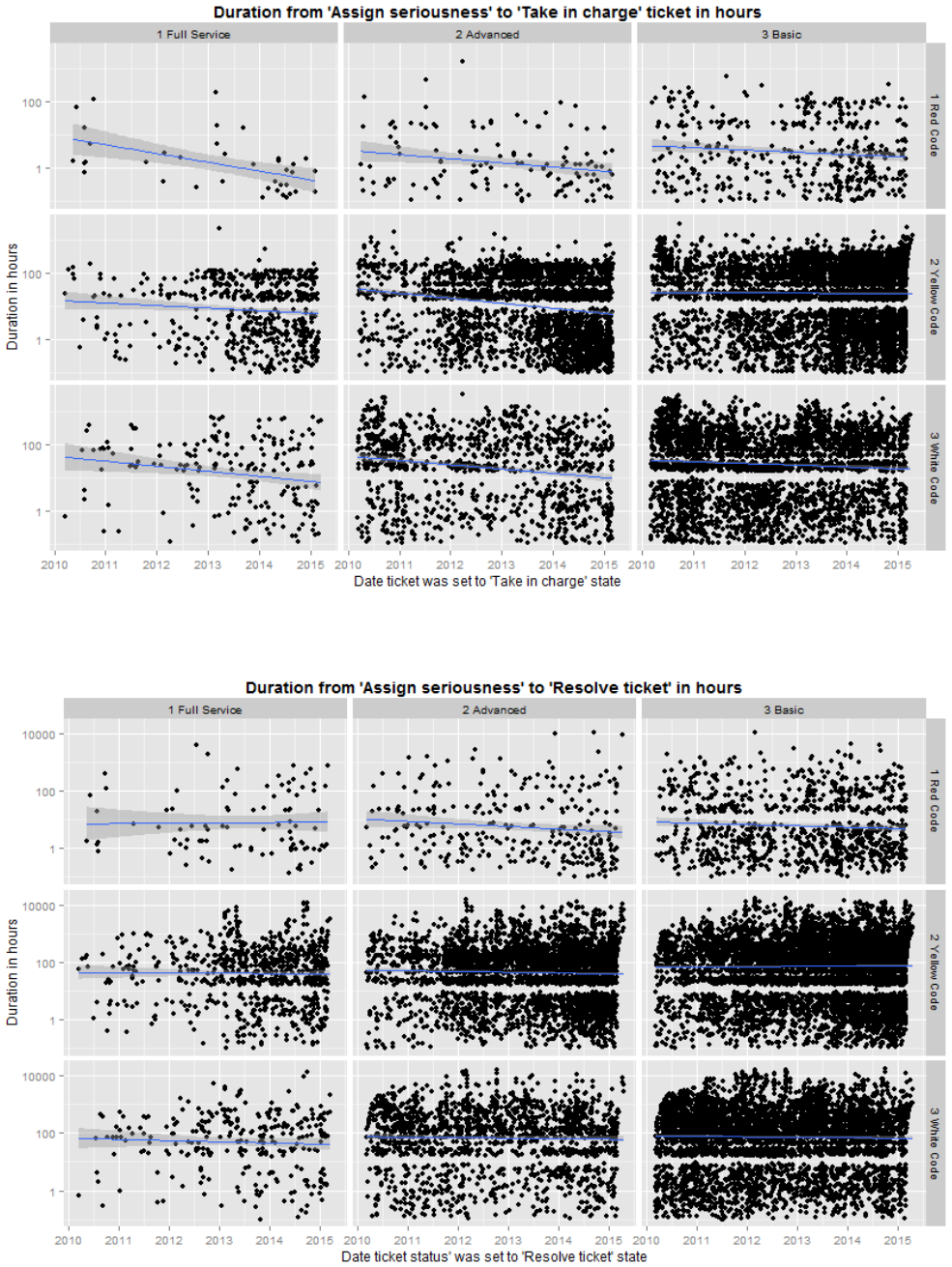
In **Table 5.1** we can see that some activities are repeated, in our analysis the first instance for each activity will be considered as the one to calculate the duration between activities. In this case, the rows highlighted in green. The second occurrence of the same activity, and the consequent ones if applicable, will be discarded from the calculations.

Withe the above considerations in place, an analysis is done to understand the durations between *Assign seriousness* to *Take in charge ticket*. In **Fig. 5.10** top diagram, a set of scatter plots showing the duration of each incident as a function of time, the y axis in logarithmic scale denotes the duration in hours. In

these plots, arranged in facets, is possible to visualize how cases (dots) preform in terms of duration between the activities in question for the respective service level and the severity of each of the cases. For instance, cases in which a customer with *Full Service* contract contacted the Help Desk team and where such cases were categorized as *Level 1 red code*, the respective duration for this facet can be found in the upper left part of the chart. Every scatter plot facet includes a trend line that helps to understand whether cases' durations are increasing, decreasing or remain constant. The trend lines help to understand whether the service level for the respective agreements and incidents severity are being met. Service level agreements rules in terms of duration are detailed in Table 3.1.

For the duration between *Assign seriousness* to *Resolve ticket* activities, **Figure 5.10** bottom diagram details how cases in the log are distributed in terms of duration, date when the case was set to *Resolve ticket* state and in service level and case severity facets. The trend line showing whether the performance in terms of meeting SLAs is of particular importance since a trend line with a negative slope means the performance is improving over time. If the trend line is positive this means that cases' durations are increasing over time.

FIGURE 5.10: Duration scatter plots by facets for *Assign severity* to *Take in charge ticket* and *Assign severity* to *Resolve ticket*.



Scatter plots are distributed depending on the severity of the case and the service level acquired by the customer. The trend line shows whether there is a positive or negative trend in the duration of cases.

CHAPTER 6

Extending the *L* Life-cycle Model* Stage 4: Operational Support

The final stage of the L^* model consists of ambitious goals like detecting deviations at run time, predicting the behaviour of a process instance or a case may have for the remaining of the process to be executed Recommend is another aspect of the stage where similarly to the prediction support, the recommendation feature of a given information system assists end users in providing a list of possible option to select from in order to make the best possible outcome of the running case[5, p. 247-257]. In this context, and during the *DMAIC Analyze* phase, several aspects of the Help Desk process were scrutinized for which it is now possible to recommend changes to our business process model with the purpose of assisting the end users of the Help Desk process in handling incidents in a more efficient and reliable way. The implementation and control of such recommendations can materialize within a framework of *DMAIC*'s *Improve* and *Control* phases which will be discussed in this chapter and that have as objectives the implementation of the new system and controlling the gains of the newly implemented improvements[20, p. 521, 585].

The final stage of the L^* *model* also considers the auditing aspect of a process where the detection of deviations from the de jure model which can assist processes owners in promoting appropriate safeguards or to propose changes to the de jure model so it aligns with the de facto model. In this context, we will discuss

how the effectiveness of audits can be improved with Process Mining tools and techniques.

6.1 Extending Process Mining *L* Life-cycle Model* Stage 4 with Baldrige Criteria™, EFQM Excellence Model™ and ISO 9001 Audits' Effectiveness

A relevant aspect for the continuous improvement of any organization that practices business improvement initiatives for excellence or that has adopted a quality management system like ISO 9001:2008 is the notion of audits, an audit is a *"Systemic evidence gathering process where audits must be independent and evidence must be evaluated objectively to determine how well audit criteria is being met"*[33]. Audits are frequently expensive to conduct, where in many occasions the lack of human and materials resources makes audits' results less effective than desired. In this context, an approach in which audits can be conducted more effectively, efficiently and reliably can be achieved, in part, with the use of Process Mining tools and techniques. Because some of the Process Mining techniques rely on historic data that can be used to conduct audits, the effectiveness and efficiency of audits can be improved considerably as auditors can also rely on this historic data to. For instance, assess process conformance, process performance and the role of human stake holders in the process. Therefore, in this chapter we will discuss how Process Mining tools and techniques using the *L* model* in its last stage can be expand their scope during audits with the aforementioned frameworks and quality management system.

6.1.1 ISO 9001:2008 Audit Requirements

ISO 9001:2008 QMS clause 5 for *Management responsibility*, in it subsection 5.6.2 *Review input* states that the input for management review shall include the results of audits, that is, the results of audits shall be reviewed by the

FIGURE 6.1: Comparing ISO and TQM Frameworks Principles[34]

Table 1. TQM and ISO 9000: 2000 principles

MBNQA Model (2001)	EFQM Model (1999)	ISO 9000 (2000)
Visionary leadership	Principle developed in the same way	Principle developed in the same way
Customer-driven excellence	Principle developed in the same way	Principle developed in the same way
Organisational and personal learning	Principle developed in the same way	Principle developed partially in the same way (focus on continuous improvement of company performance)
Valuing employees and partners	Principle developed in the same way	Principle developed partially in the same way (focus on both personnel and supplier involvement)
Agility	Principle not developed	Principle not developed
Focus on the future	Principle not developed	Principle not developed
Managing for innovation	Principle given less emphasis	Principle not developed
Management by facts	Principle developed in a similar way, but with more emphasis on processes	Principle developed in a similar way, but with more emphasis on processes and on their interconnections
Public responsibility and citizenship	Principle developed in the same way	Principle not developed
Focus on results and creating value	Principle developed in the same way	Principle developed in the same way

leadership at planned intervals to guarantee continuing suitability, adequacy and effectiveness [9, p. 5].

Similarly, clause 8 *Measurement, analysis, and improvement* details in its subsection 8.2.2 *Internal audit* that *"the organization shall conduct internal audits at planned intervals to determine whether the quality management system conforms to planned arrangements and whether the quality management system is implemented and maintained"*[9, p. 12]. The *Internal audit* subsection clause also states that *"an audit program shall be planned taking into account the relevance and status of the processes and areas to be audited as well as previous audits results"*. In this context the norm also states that *"the audit criteria, scope, frequency and methods shall be defined and that the selection of auditors and the execution of audits shall ensure objectivity and impartiality of the audit process"*. Finally, the clause states that *"auditors should not audit their own work for obvious reasons, and that there must be a documented procedure to define responsibilities and requirements for the planning and conducting of audit"*[9,

p. 12]. ISO 9001:2008 QMS also states that *"for the continual improvement of the quality management system the organization shall continually improve the system through the use of audits' results"*[9, . 14].

The above mentioned criterion can be used to expand the final stage of the *L* model* as it can give broader scope to project managers in charge of Process Mining projects, in particular if the project has as one of its goals the auditing of a process.

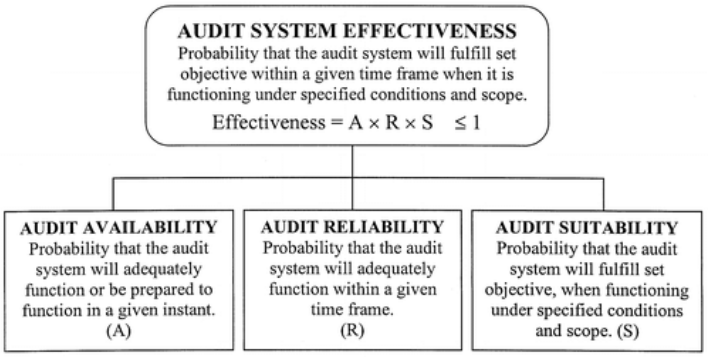
6.1.2 Baldrige Criteria™ and EFQM Excellence Model™ Process Improvement through Audits

The EFQM Excellence Model™ details in its *Processes, Products & Services* enabler in its sub-sections 5.a and 5.d the following criterion[13, p. 16]:

- *"The use of a framework of key processes to implement the organization strategy."*
- *"Manage processes end to end, including processes that extend beyond the boundaries of the organization."*
- *"Ensure process owners understand their role and responsibility in developing, maintaining and improving processes."*
- *"Develop a meaningful mix of of process performance indicators and related outcome measures that enable the review of the efficiency and effectiveness of the key processes and their contributors towards the strategic goals."*
- *"Use data on the current performance and capabilities of the processes."*
- *"The organization shall develop an effective and efficient value chain to ensure the promised value proposition can be adequately delivered."*
- *"It must be ensured that people have the necessary resources, competences and empowerment to maximize the customer experience."*

The Baldrige Criteria™ in its criteria for performance excellence details in section number 4 *Measurement, Analysis, and Knowledge Management* and section

FIGURE 6.2: Audit System Effectiveness Measures



Process Mining projects via *L* model* can contribute to increase the probability for the audit system to meet the objectives[35, p. 691].

b. *Performance Analysis and Review* that the leadership must consider the review of the organization's performance and capabilities[17, p. 17]. In section c. *Performance Improvement* is detailed that one of the criteria is for leadership to use findings from performance reviews –see 4.1 b already stated– to develop priorities for continuous improvement and opportunities. Additionally, Baldrige Criteria™, in its *Measurement, Analysis, and Knowledge Management* criteria and in its 4.2 section *Knowledge Management, Information, and Information Technology* lists that leadership must consider the organizational knowledge in terms of:

- Collecting and transferring knowledge within the workforce
- Blending and correlating data from different sources to build new knowledge

Within the same section criteria, the program also states on the need to make data and information available[17, p. 18]. Here is worth mentioning that the Baldrige Criteria™ emphasizes more on a criteria to make information available and to transfer knowledge within the organization, this a relevant criteria that the EFQM Excellence Model™ does not discuss in such detail.

Given the fact that Baldrige Criteria™ and EFQM Excellence Model™ frameworks do not necessarily mention the concept of audits, in practice that organization needs to conduct a process of systemic gathering of evidence that can be

used in assisting leadership in identifying opportunity areas for the continuous improvement. For this reason is relevant to include such frameworks in the discussion regarding how Process Mining using the *L* model* can leverage from such criteria to understand in what extent the criterion detailed in this subsection can in practice expand the model in its last stage.

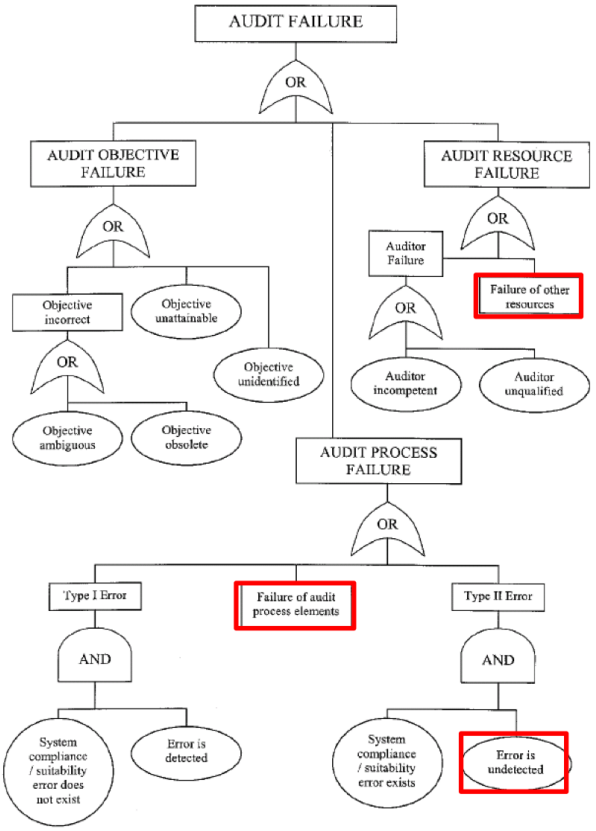
6.1.3 Improving Audit Availability, Reliability, and Suitability

As it has been discussed above, where a brief comparison for the *Management by facts* criteria is detailed in **Figure 6.1**, it is a good practice for the competitive advantage of an organization to gather objective, impartial, and easily available information that can be transferred within the organization's workforce and of course for such information to be made available for management for review in order to make important decisions. A key aspect of such information gathering, as clearly outlined by ISO 9001:2008 QMS, is the collection of information through audits. As any practice subject to improvement, audits have effectiveness limitations which are limited by three aspects:

- Audit availability
- Audit reliability
- Audit suitability

The three measures detailed above, which are illustrated in **Figure 6.2**, are exposed to a number of factors that can cause audits to fail, these factors can manifest in the *Reliability* aspect of the audit as a lack of sufficient amount of evidence which can cause an audit error or by deficiencies in material resources and lack of available time which can cause damage to the audit's reliability. **Figure 6.3** details a fault tree diagram for audits, in this diagram three areas can benefit from Process Mining:

FIGURE 6.3: Fault Tree Diagram for an Audit[35, p. 693]

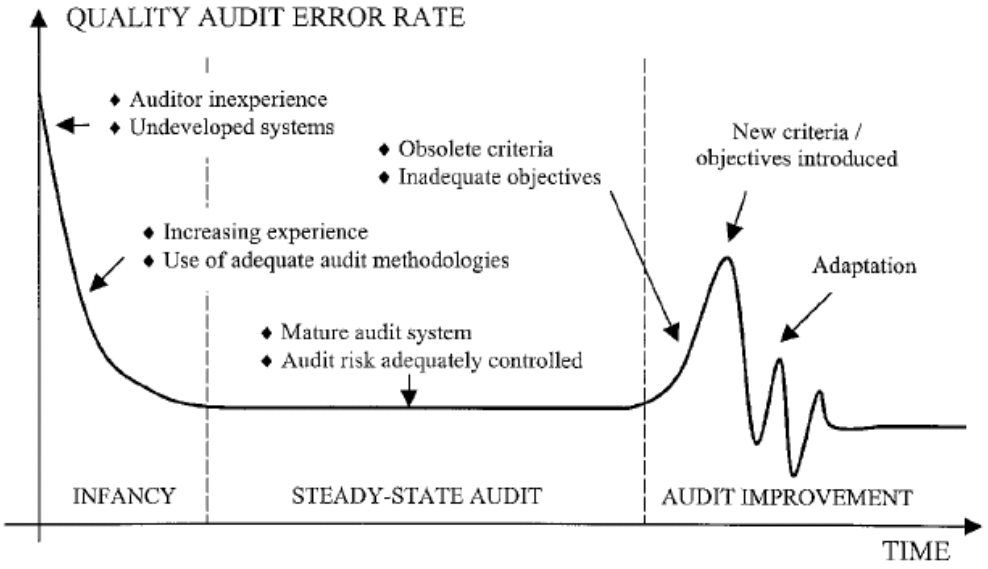


The *L* model* for Process Mining projects in its stage 4 for operational support can contribute to decrease error sources outlined in red.

- A Type II Error can be mitigated by making the detection of errors more likely, because Process Mining tools and techniques, in particular stage 4 *Operational support* aims at checking for deviations manifested in *post-mortem* data in event logs[1, p. 244].
- Sine Process Mining relies on persisted data, the failure of audit process elements and the failure of other resources can be mitigated given that persisted data can be made readily available by auditors request before of after an audit is conducted.

Process Mining tools and techniques through the *L* model* can bring a relevant set of tools for auditors to conduct their work, in this context, better tools can lead to the improvement of adequate audit methodologies and assisting auditors in mitigating inadequate objectives with the use for example of process mapping

FIGURE 6.4: The Bathtub Curve in Auditing



Process Mining projects via *L* model* can contribute to decrease the audit error rate during the infancy stage by improving the use of adequate audit methodologies[35, p. 694].

and conformance checking which are more efficient and intuitive ways to audit processes compared to auditing processes from a silo type approach or auditing processes in a time frame limited to that when the actual audit takes place.

The bathtub curve depicted in **Figure 6.4** explains the evolution of audits. during the infancy phase of the curve the audit error rate is considerably higher, in this stage Process Mining can contribute to mitigate the error rate by assisting auditors in determining adequate audit methodologies because Process Mining *L* model* in its final stage uses conformance checking or organizational mapping to be used as a auditing tool. For an auditor to grasp Process Mining tools and techniques and getting used to the *L* model* could represent a considerable steep learning curve, but once these skills are gained the tools and techniques provided by Process Mining can be a powerful asset for auditors to conduct their work.

Auditing process with Process Mining tools and techniques can then bring considerable value to auditors in terms of efficiency, reliability, accuracy, and availability, all of which can mitigate audit errors in a considerable measure.

6.2 Extending Process Mining L* *Life-cycle Model* Stage 4 with Six Sigma's *Improve* and *Control* Phases

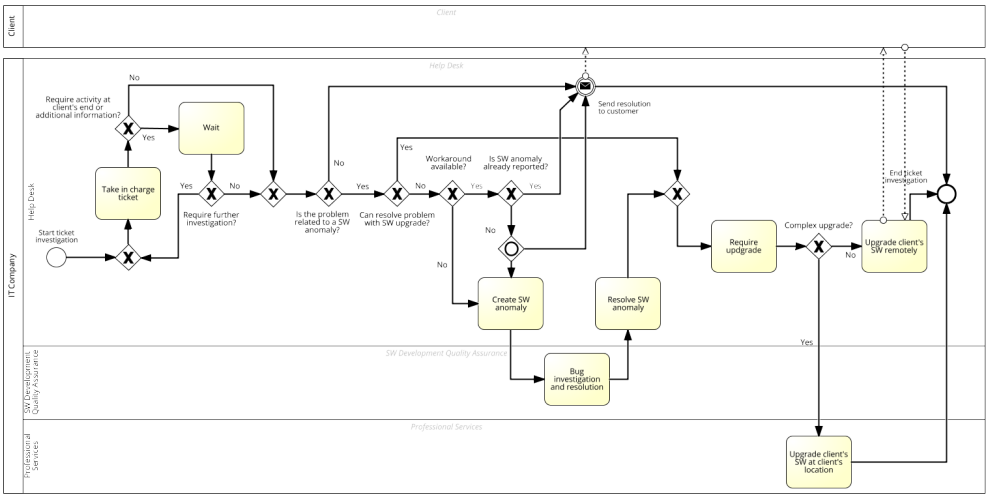
This section of the chapter will focus in detailing the proposed improvements to the system and outlining a plan to maintaining the benefits of such improvements. At the time this thesis was concluded the statistical validation of the improvements, which is part of *DMAIC*'s *Control* phase, was not conducted, however, the chapter will discuss how the improvements made can be controlled in the context of a business process improvement.

6.2.1 Improving the Help Desk Ticket Investigation Sub-process

In section **5.2.1** it was discussed that after running conformance checking diagnostics activity *Resolve SW anomaly* fitness measure had considerably lower values compared to the rest of the activities in the process, because of this, the project team decided to qualify this finding as a candidate for *DMAIC Improve* and *Control* phase which in practice will serve as the extension of stage 4 of the *L* model*. The model in its final stage intends to offer the results gained in previous stages of the model to be delivered to end users, in this context, is possible to determine that after analyzing the Help Desk process from a conformance checking point of view the project team can recommend improvements that can serve the users of the Help Desk process where such improvements can be maintained by continuously analyzing the ongoing process instances to ensure desirable conformance checking values.

Given that *DMAIC*'s *Improve* phase aims at lying down a set of proposals, the selection of such proposal can take place with Pugh Concept Selection Method[20, 521] or business process simulation tools[36] in order to asses how the new business process model performs.

FIGURE 6.5: Improved Ticket Investigation Sub-process



The ticket investigation sub-process has been improved by re-configuring and adding safeguards for incidents that involve potential SW anomalies.

The improvement proposals for the Help desk process are detailed in **Figure 6.5**, these proposals consist of adding additional safeguards, modeled as exclusive gateways and an additional inclusive gateways. The goal ask additional questions to the resources dealing with Help Desk incidents that are related to potential SW anomalies, the first exclusive gateway where a resource first must decide whether a work around for the SW anomaly is available, in case a workaround is available then the resource working in the case decides if such SW anomaly has been reported already. If the anomaly has been reported or not the Help Desk team member sends the resolution to the customer and may decide whether to log a potential SW anomaly or not. For this, an inclusive gateway has been modeled to hint that executing *Create SW anomaly* activity is not mandatory and that the Help Desk incident can be closed without the need to have activity *Resolve SW anomaly* executed.

In case there is no workaround available for the potential SW anomaly, the Help Desk resource executes *Create SW anomaly* activity, then the process waits for such SW anomaly to be resolved to execute *Resolve SW anomaly* activity, such as it is modeled in the business process detailed in section **4.3.2**.

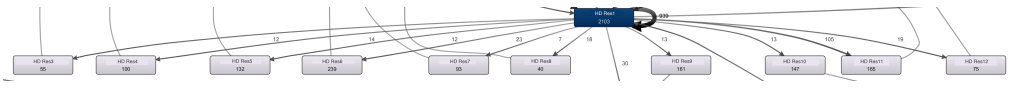
The improvements detailed in the aforementioned paragraphs can then represent

an business process implementation in which the recommendations in the form of gateways can assist the end users of the process to better handle incidents so they comply with the de jure model. The recommendation feature in an information system to better handle process instances or incidents is one of the goals in the *Operational support* stage of the *L* model*[5, p. 257-258].

6.2.2 Organizational Improvement

An other opportunity identified in the *Measure* phase, see appendix in **section A.7**, is how Help Desk incidents involving potential SW anomalies are handled among stake holders in the process. A proposed improvement is to assign incidents involving potentials SW anomalies to one or two Help Desk team members so the variation in terms of knowledge for these cases is reduced, that is, a new role where a Help Desk team member specializes in incidents involving SW anomalies can be created. So the team member with this profile can manage such complex cases together with the SW Quality Assurance team instead of multiple Help Desk team members having to handle such incidents with the SW Quality Assurance team. **Figure 6.6** details how multiple resources are handling incidents involving SW anomalies, the goal in mind is to assign such cases to one or two team members that have specialist roles in handling incidents with SW anomalies.

FIGURE 6.6: Handing-over for SW Anomaly Incidents



Incidents involving SW anomalies are being handed over from the manager to other Help Desk team members, the improvement proposal is to create a specialist position that only handles SW anomaly incidents so this specialist can interface more effectively with the SW Quality Assurance team

6.2.3 Maintaining Gains for the Improved Business Process

DMAIC's *Control* phase serves the purpose of statistically validating that the new process meets the objectives and benefits outlined through the project. There is also the challenge to institutionalize the the new changes in the process, this task is often overlooked as it is assumed that the stakeholders in the process will somehow be briefed about the newly implemented changes and therefore start using the new processes flawlessly. This assumption could create confusion and conflict among team members.

In order to prevent such drawbacks, a set of methods can be followed to ensure that those involved in the new process are properly briefed and involved so the implemented improvements are maintained[20, p. 586]:

- **Update policy changes** within the organization responsible for the new process.
- **Update work procedures** describing the new changes in the process.
- **Modify audit criteria** so there is a criteria in place for auditors to consider in case the new process is part of an audit
- **Update budgets and financial goals**, because the new process means a more efficient operation (otherwise the Six Sigma project would be justified) budgets and finical goals must be modify accordingly so the improvements in the new process are reflected in financial statements.
- **Update training methods and materials** so these include the improvements made to the new process.
- **Update information systems** so the improvements implemented in the new process are in sync with automated tasks and other automated support processes.

Aside from the aforementioned ways to maintain the gains for the new process, it is important to consider whether [21, p. 342] the measures in place before the

new process was implemented are still necessary, if the new process has been implemented with goals where some of the activities in the process are executed more efficiently, then it must be revised whether the measures remaining in the new process are still necessary.

Another important aspect is to understand whether the new process is actually improving internal or external customer satisfaction[21, p. 342], this can be achieved in part, via voice of the customer (VOC)[37] through surveys or other methods that can quantify whether the implemented and maintained improvements for the new process have the desired impact for customers.

CHAPTER 7

Contributions, Limitations, and Further Work

This thesis has examined the extension of the *L* life-cycle model* for Process Mining projects with business improvement frameworks such as the Baldrige Criteria for Performance Excellence for Business and Nonprofit^{TM} and the European Foundation for Quality Management Excellence Model^{TM}, as well as ISO 9001:2008 Quality Management System. This work also postulates how the *L* model* in every one of its stages can be expanded with Six Sigma's *Define, Measure, Analyze Improve, and Control (DMAIC)* model methodology and how the extended *L* model* in the basis of a Process Mining project has been used in guiding an effort for the improvement of a Help Desk and Software Quality Assurance processes of an Italian IT Company.

7.1 Contributions

As stated in the introduction section of this thesis, the *L* life-cycle model* lacks a framework that includes goals and criteria set forth by organizations leadership as well as a framework that properly articulates internal and / or external customers, the boundaries and targets of the project during its inception. As well, the *L* model* shortfalls in identifying the need for a project to analyze causes of

variation, so such causes are prioritized in order to leverage from them in order to make and maintain improvements for the process being studied.

In view of the shortfalls stated in the aforementioned paragraph, the work in this thesis has demonstrated that the *L* life-cycle* model can in practice be expanded by adequately linking it to criteria governed by business improvement frameworks. In particular, chapter 3 through 6 demonstrate how in the basis of the Baldrige Criteria^{TM}, the EFQM Excellence Model^{TM}, and ISO 9001:2008 QMS, a Process Mining project can be justified and expanded due to the principles of *managing by fact* and the need of managing and improving business processes in a continuous fashion.

In terms of contributions by extending the *L* model* with Six Sigma's *DMAIC* model, chapter 3 discusses the extension of *L* model plan* stage with *DMAIC*'s *Define* phase where a project charter, customer requirements like service level agreement (SLA), and a top view of the business processes in the IT Company were introduced.

The contributions made in chapter 4 demonstrate that the *extract* stage from the *L* model* can successfully be extended with *DMAIC*'s *Measure* phase where an important achievement is the articulation of how an existing system can be measured in the context of a process mining project and how such measuring system can be used to align the stakeholders in the project in preparation for subsequent stages of the project.

In chapter 5 the extension of stages 2 and 3 of the *L* model* with *DMAIC*'s *Analyze* phase contributes in the sense of identifying causes of variation in the several perspectives of the Process Mining project, a particularly relevant contribution is the identification of activities that systematically present lower conformance checking values and how these variations contribute in diminishing the processes overall conformance checking performance. Another contribution is how releases of new products can be measured in a time perspective to understand how the performance of several releases compares against each other, another approach to identify causes of variation.

The extension of *L* model* stage 4 with *DMAIC*'s *Improve* and *Control* phases is discussed in chapter 6, here contributions were made in terms of demonstrating that once opportunities areas are identified, these can be prioritized and implemented accordingly as well as how these can be maintained. The implementation and maintenance of improvements made to business processes is an important aspect of the project, therefore the importance of the contribution mentioned earlier in the context of a Process Mining improvement effort.

In consideration of the contributions discussed in the previous paragraphs, it is important to highlight that in the measure that improvement projects are properly articulated with leadership's sponsorship, involvement of processes' stakeholders and domain experts projects are better positioned to commence with a higher degree of commitment by those involved. Similarly, by having clearly identified and communicated goals, such projects are more likely to succeed in identifying, implementing and maintaining improvements. That is, Process Mining projects using the extended *L* life-cycle model* with business improvement frameworks like the Baldrige Criteria for Performance Excellence^{TM} and the EFQM Excellence Model^{TM} as well as ISO 9001:2008 Quality Management System and Six Sigma's *DMAIC* model could, in some measure as discussed in previous chapters, improve the prospect of such projects in making a greater impact in organizations' efforts to improve business processes.

7.2 Limitations

The work in this thesis fell short in including the verification of improvements made to the Help Desk process in the Italian IT Company, this was mainly due to insufficient time so the new behaviour of the improved process could be measured and verified. The work in this thesis also has limitations in regards to the business improvement frameworks criteria in the context that extending the *L* Process Mining model* with such criteria could have been conducted in an organization that actually practices one of the business improvement frameworks discussed in previous chapters.

An additional limitation is related to the fact that the project did not involve a business case where financial benefits to the business after the completion of the project could have been identified and documented in the thesis. Similarly, this thesis did not work in the context of institutionalizing the extended *L* model* for Process Mining projects in an organization where such extended model could contribute in the organization's improvement efforts.

7.3 Further Work

As discussed in section 5.2.1, conformance checking diagnostics analyses were made to the Italian IT Company Help Desk process, during the analysis phase for this task it was realized that the available Process Mining software that can perform conformance checking diagnostics results in BPMN 2.0 language is very limited. Currently the only available Process Mining software solution that can do this is University of Padua's plug in that performs *Conformance Checking of a Petri Net with results in BPMN* with ProM 6 software. Such plug in has substantial limitations. For instance, to be able to get conformance checking diagnostics results in BPMN 2.0 language it was necessary to first create a simplified business process diagram in BPMN 2.0 language, then this had to be matched with a Petri net that would reproduce the same behaviour for which testing had to be made in order to verify that the Petri net would produce the desired behaviour. These two tasks took a considerable amount of time, about two to three business days. Given that the process in question only involves 9 activities and 11 exclusive gateways, it would be cumbersome to attempt the same task with a much more complex process.

Therefore, there is the need to develop Process Mining software solutions that can deliver conformance checking results in BPMN 2.0 language more efficiently. Getting conformance checking results in BPMN 2.0 language is an important feature in Process Mining software given that such language has been widely adopted within the Business Process Management domain.

Bibliography

[1] W. M. P. van der Aalst, M. Dumas, F. Maggi, and et al, "Process Mining Manifesto," 2011. [Online]. Available: http://www.win.tue.nl/ieeetfpm/lib/exe/fetch.php?media=shared:process_mining_manifesto-small.pdf

[2] W. M. P. van der Aalst and A. J. M. M. Weijters, "Process mining: a research agenda," *Computers in Industry - Special issue: Process/workflow mining*, vol. 53, no. 3, pp. 231–244, April 2004. [Online]. Available: http://dx.doi.org/10.1016/j.compind.2003.10.001

[3] A. Rozinat, "How is process mining different from..." 2014. [Online]. Available: http://fluxicon.com/blog/2014/02/how-is-process-mining-different-from/

[4] W. M. P. van der Aalst and et al, "Business process mining: An industrial application," *Information Systems*, vol. 32, no. 5, pp. 713–732, July 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306437906000305

[5] W. M. P. van der Aalst, *Process Mining - Discovery, Conformance and Enhancement of Business Processes*. Springer, 2011. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-19345-3

[6] M. Bozkaya, J. Gabriels, and J. M. E. M. van der Werf, "Process diagnostics: A method based on process mining," in *International Conference on Information, Process, and Knowledge Management, eKNOW 2009, Cancun, Mexico, February 1-7, 2009*, 2009, pp. 22–27. [Online]. Available: http://dx.doi.org/10.1109/eKNOW.2009.29

[7] M. L. van Eck, X. Lu, S. J. J. Leemans, and W. M. P. van der Aalst, "PM ˆ2 : A process mining project methodology," in *Advanced Information Systems Engineering - 27th International Conference, CAiSE 2015, Stockholm, Sweden, June 8-12, 2015, Proceedings*, 2015, pp. 297–313. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-19069-3_19

[8] "Understanding process behaviours in a large insurance company in australia: A case study," in *Advanced Information Systems Engineering*, ser. Lecture Notes in Computer Science, C. Salinesi, M. Norrie, and . Pastor, Eds., 2013, vol. 7908. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-38709-8_29,publisher={SpringerBerlinHeidelberg},keywords={processmining;casestudy;businessprocessmanagement},author={Suriadi, SuriadiandWynn,MoeT.andOuyang,ChunandterHofstede,ArthurH.M. andvanDijk,NienkeJ.},pages={449-464},language={English}

[9] "The ISO 9001:2008 Standard," http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=46486, accessed: 2015-06-12.

[10] "The Baldrige Program," http://www.nist.gov/baldrige/about/index.cfm, accessed: 2015-06-11.

[11] "The European Foundation for Quality Management Excellence Model," http://www.efqm.org/efqm-model/efqm-model-in-action-0, accessed: 2015-06-22.

[12] "Disco software by Fluxcion," https://fluxicon.com/disco/, accessed: 2015-08-08.

[13] E. F. for Quality Management, *EFQM Excellence Model*. EFQM Afnor, 2012.

[14] "Porter Model: The Five Competitive Forces That Shape Strategy," https://hbr.org/2008/01/the-five-competitive-forces-that-shape-strategy, accessed: 2015-07-09.

[15] "EFQM Excellence Model, PROCESSES, PRODUCTS & SERVICES," http://asq.org/learn-about-quality/new-management-planning-tools/overview/relations-diagram.html, accessed: 2015-07-09.

[16] "The Baldrige Program, Who we are." http://www.nist.gov/baldrige/about/index.cfm, accessed: 2015-07-07.

[17] "Quality Texas 2015-2016 Baldrige Excellence Framework for Business and Nonprofit," http://quality-texas.org/wp-content/uploads/2014/11/2015-2016-Baldrige-Excellence-Framework_B-NP_Examiner_Use_Only.pdf, accessed: 2015-08-06.

[18] "Baldrige National Quality Program, CEO Issue Sheet: BALDRIGE, SIX SIGMA, & ISO: Understanding Your Options." http://www.nist.gov/baldrige/publications/upload/Issue_Sheet_SS.pdf, accessed: 2015-07-07.

[19] "The ISO Survey 2013," http://www.iso.org/iso/iso_survey_executive-summary.pdf?v2013, accessed: 2015-06-12.

[20] P. Keller and T. Pyzdek, *The Six Sigma Handbook, Fourth Edition*. McGraw-Hill Education, 2014.

[21] P. Harmon, "Business process change," in *Business Process Change (Second Edition)*, second edition ed., ser. The MK/OMG Press, P. Harmon, Ed. Burlington: Morgan Kaufmann, 2007.

[22] "About The Baldrige Excellence Framework Criteria," http://www.nist.gov/baldrige/publications/bus_about.cfm, accessed: 2015-07-07.

[23] "David A. Garvin: How the Baldrige Award Really Works," https://hbr.org/1991/11/how-the-baldrige-award-really-works, accessed: 2015-07-07.

[24] "About Bugzilla." https://www.bugzilla.org/about/.

[25] "Supply Chain KPI Examples." http://www.klipfolio.com/resources/kpi-examples/supply-chain, accessed: 2015-08-13.

[26] I. Rosam and R. Peddle, *Creating a Process-based Management System for ISO 9001:2000 and Business Improvement*. BSI British Standards Institution, 9 2009. [Online]. Available: http://amazon.com/o/ASIN/0580676579/

[27] "ProM Tools Official Website," http://www.promtools.org/doku.php, accessed: 2015-08-05.

[28] "Life Cycle of a Bug, The Bugzilla Guide 2.18.6 Release." https://www.bugzilla.org/docs/2.18/html/lifecycle.html, accessed: 2015-07-07.

[29] "Disco user-guide." http://fluxicon.com/disco/files/Disco-User-Guide.pdf, accessed: 2015-08-09.

[30] "ProM 6 Documentation: Replay Log on Petri Net for Conformance Analysis." http://blog.adriansyah.info/archives/417, accessed: 2015-07-28.

[31] "University of Padova PROMPT Project." http://www.unipd.it/international-highlights/node/358, accessed: 2015-07-28.

[32] "Tableau Business Intelligence Software Official Website." http://www.tableau.com, accessed: 2015-07-29.

[33] "Praxiom ISO 9000 2105 definitions," http://www.praxiom.com/iso-definition.htm#Audit, accessed: 2015-08-06.

[34] S. Biazzo, "The new iso 9001 and the problem of ceremonial conformity: How have audit methods evolved?" *Total Quality Management & Business Excellence*, vol. 16, no. 3, pp. 381–399, 2005. [Online]. Available: http://dx.doi.org/10.1080/14783360500054145

[35] S. Karapetrovic and W. Willborn, "Quality assurance and effectiveness of audit systems," *International Journal of Quality & Reliability Management*, vol. 17, no. 6, pp. 679–703, 2000. [Online]. Available: http://dx.doi.org/10.1108/02656710010315256

[36] "Business Process Simulation Tool Survey." http://www.researchgate.net/publication/228358579_Business_process_simulationa_tool_survey, accessed: 2015-08-09.

[37] "American Society for Quality on Voice of the Customer (VOC) ," http://asq.org/service/body-of-knowledge/tools-voice-of-the-customer, accessed: 2015-06-22.

Appendices

A.1 Appendix A

D	**Define** the goals of the improvement activity.
M	**Measure** the existing system.
A	**Analyze** the system to identify ways to eliminate the gap between the current performance of the system or process and the desired goal.
I	**Improve** the system.
C	**Control** the new system.

The Six Sigma DMAIC approach [20, p. 4].

A.2 Appendix B

Project Charter	
Project Name: Help Desk and Bugzilla Business Processes Mapping and Analysis	
Opportunity/Problem Statement: 1. There is a business need to build a de jure model taking in consideration event logs as well as stakeholders input via interviews and in what degree the de jure model is being followed by those involved in the help desk and Bugzilla processes. 2. There is a need to understand how the stakeholders are interacting in the Help Desk and Bugzilla processes. 3. There is a need to understand in what degree the service level agreement (SLA) metrics are being followed. 4. There is a need to understand what are the different data attributes related to help desk incidents and how these impact the SLA. 5. There is a need to understand how the different software releases are influencing the number of Help Desk incidents related to potential SW anomalies issues. 6. There is a need to understand if the number of Help Desk incidents is increasing.	
Business Case/Impact: An understanding on how the process is being executed and how this compares to a de jure model will serve the purpose in identifying potential deviations and efficiencies in the process so management can take action in addressing these potential non conformances. As well, an understanding on how the SLA is being met can help the stakeholders in the process to identify what can be done to improve such metrics and what are the incidents characteristics that may be influencing the throughput time of incidents.	
Team Members: -Business Process Analyst Intern (Agustin Guerrero) -Process Mining R&D Engineer -Social Network Analysis SW Developer -SW Quality Assurance Engineer -Help Desk Team Manager -Help Desk Team Member 1 -Help Desk Team Member 2 -Professional Services Manager	**Project Owner:** J. Agustin Guerrero T. **Project Sponsor:** Software Development Manager
Project Performance Indicator(s): The project indicators will be measured in terms of completion for the project several stages based in the following deadlines: -Week 1: Selection of data to be extracted which must include attribute data that can help in determining SLA metrics. -Week 2 to week 4: Data extraction and creation of log. -Week 5 to week 7: Business process mapping and de jure model creation. -Week 8 to 9: Social network discovery and analysis for Help Desk and Bugzilla processes. -Week 10 to week 12: Incidents data attribute analysis. -Week 13 to 16: Recommendations on business process improvements and actions to control the new business process after recommendations are implemented.	

A project charter is essential in the *Define* phase of the Six Sigma DMAIC approach. [20, p. 245-247].

A.3 Appendix C

Table 7.1 Suitable and unsuitable system KPIs

Process	Suitable system KPI	Unsuitable system KPI
Understanding the market	Percentage of market reviews required in a process fully presented, validated and accepted by its planned date	Customer satisfaction rating
Business planning	Business plan agreed by board and group by required date	Number of initiatives in the business plan
Developing products	Percentage of new products that successfully deliver target profit	Number of new products released
Winning business	Value of sales made that have been confirmed by rest of business as achievable divided by sales target	Value of sales achieved
Delivering products	Customer satisfaction ratio against target	Deliveries made on time
Measuring performance	Percentage of accurate performance reports produced to timetable	Number of reports delivered
Making improvements	Percentage of planned deliverables achieved through major improvement projects	Number of improvement projects initiated
Managing people	Staff morale rating against target	Percentage of staff undergoing training
Managing assets	Percentage availability against plan	Number of maintenance schedules achieved
Managing finance	Profit achieved against target	Average debtor days

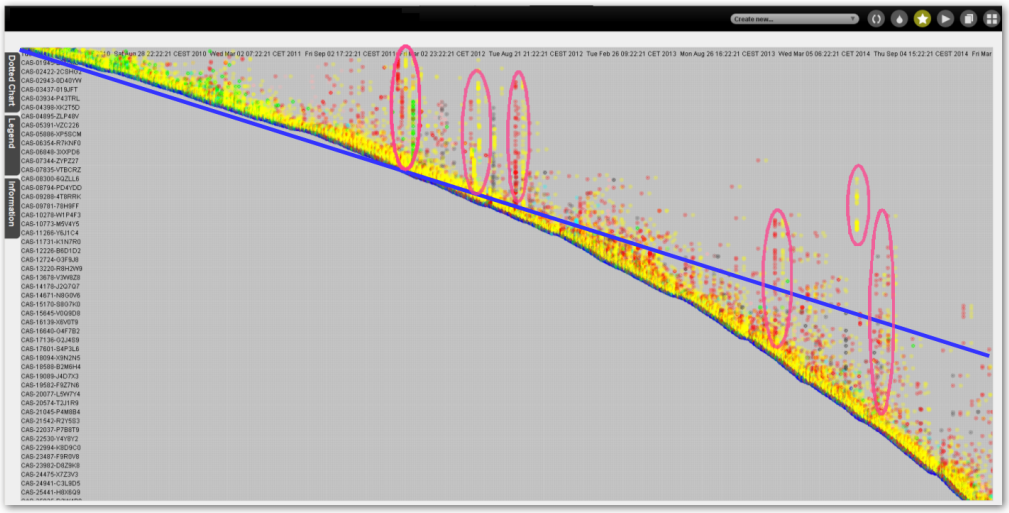
Example of suitable and unsuitable system KPIs [26, p. 56].

A.4 Appendix D

Filed Name	**Data Type**	**Source**
Case ID	String	CRM
Activity	String	CRM
Resource	String	CRM
Complete Timestamp	Date Time	CRM
OrigTraceID	Alphanumeric	Bugzilla
added	String	Bugzilla
assignedto	String	Bugzilla / CRM
bug_id	Numeric	Bugzilla
bug_severity	String	Bugzilla
cf_customer	String	Bugzilla
cf_rifcrm	String	CRM
customer	String	CRM
description	String	CRM
eventid	String	CRM
incidentname	String	CRM
lctrans	String	Bugzilla
owner	String	Bugzilla
priority	String	Bugzilla
product	String	CRM
seriousness	String	CRM
servicelevel	String	CRM
servicetype	String	CRM
version	String	Bugzilla / CRM
workgroup	String	CRM

Data types extracted from Bugzilla and Microsoft Dynamics CRM systems.

A.5 Appendix E



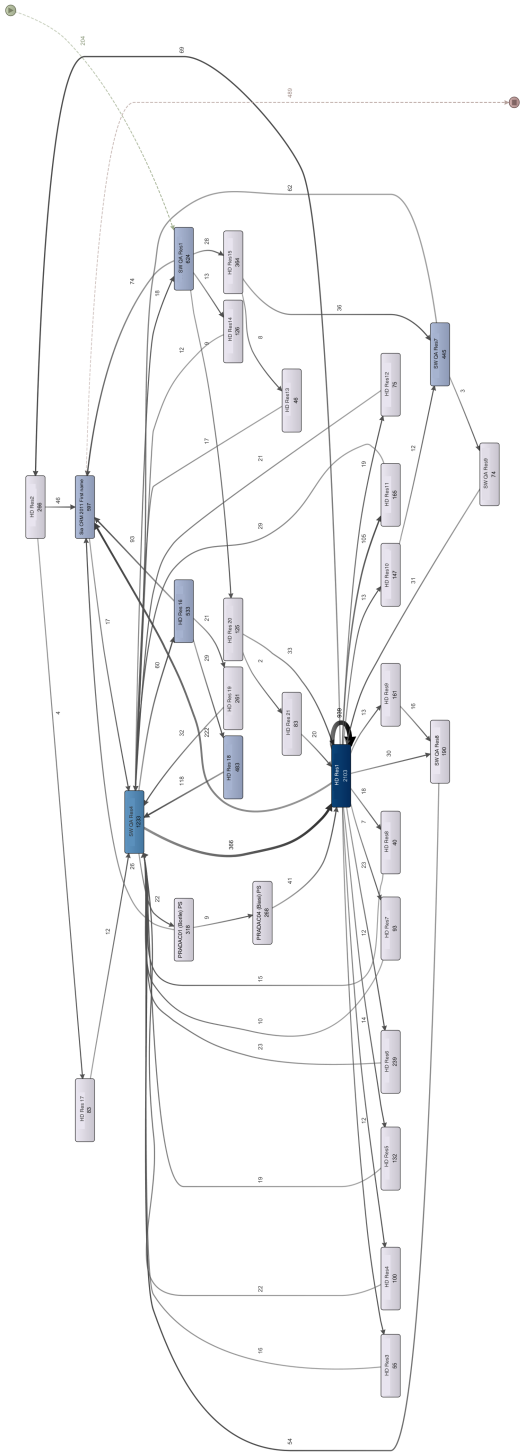
In this XDotted chart it can be seen that the arrival of cases is increasing over time, given the slope indicated by the blue line, the arrival rate is steeper than the blue line. It ca also be noted that there are activities that are being executed in batches, this is indicated by the activities noted inside the pink ovals.

A.6 Appendix F

Fitness Measure	Fitness Value	Date Period
Overall process	0.9562	2010 First Half
Insert ticket	0.9529	2010 First Half
Assign seriousness	0.8315	2010 First Half
Take in charge ticket	0.9278	2010 First Half
Wait	0.9490	2010 First Half
Create SW anomaly	0.9194	2010 First Half
Resolve SW anomaly	0.0702	2010 First Half
Require upgrade	0.0000	2010 First Half
Resolve ticket	1.0000	2010 First Half
Closed	0.9843	2010 First Half
Overall process	0.9606	2010 Second Half
Insert ticket	0.9907	2010 Second Half
Assign seriousness	0.8449	2010 Second Half
Take in charge ticket	0.9432	2010 Second Half
Wait	0.9502	2010 Second Half
Resolve SW anomaly	0.3165	2014 Second Half
Require upgrade	0.7913	2014 Second Half
Resolve ticket	0.9700	2014 Second Half
Closed	0.9775	2014 Second Half
Overall process	0.9492	2015 First Half
Insert ticket	1	2015 First Half
Assign seriousness	0.8209	2015 First Half
Take in charge ticket	0.9426	2015 First Half
Wait	0.9575	2015 First Half
Create SW anomaly	0.8776	2015 First Half
Resolve SW anomaly	0.3385	2015 First Half
Require upgrade	0.8148	2015 First Half
Resolve ticket	0.9609	2015 First Half
Closed	0.9548	2015 First Half

Data structure used to perform conformance checking analyses for the Help Desk process, the data structure design aims at performing time series analysis and summary statistics.

A.7 Appendix G



Organizational map of Help Desk Process and Software Quality Assurance resources, when there are incidents involving potential SW anomalies, cases are passed among team members.

A.8 License

Non-exclusive licence to reproduce thesis and make thesis public

I, Jesús Agustín Guerrero Topete (date of birth: 12th of November 1980),

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

1.1 reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

1.2 make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

Extending the L^* Process Mining Model with Quality Management and Business Improvement Tools and Techniques

supervised by Dr. Marlon Dumas

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, 13.08.2015