# Internship Report:

## Working As Data Analyst for Umbian Inc.

By

Diaa Rafiq

B00560513

Diaa.rafiq@dal.ca

Performed at

Umbian Inc.,

38 Solutions Drive, Suite 300
Halifax, Nova Scotia B3S 0H1

In partial fulfillment of the requirements of the Master of Health Informatics Program,

Dalhousie University

Report of Internship for the period June 2 – September 4, 2015

Date Submitted: September 23, 2015

# Acknowledgment

This report has been written by me and has not received any previous academic credit at this or any other institution.

I would like to thank Mr. Caleb Woodford for the support and guidance he provided as my work supervisor at Umbian Inc.

Diaa Rafiq

# Contents

# Executive summary

Umbian Inc. is a health solution software company that have two main products for managing sleep apnea therapy. The first product is U-sleep which uses 'Manage by Exception' to manage a large amount patients. The second product is MyAir website, which is a patient oriented website that gives patients scores based on device use and therapy compliance. Working as a data analyst at the company has allowed me to utilize my SQL, web developing and statistical skills to achieve the tasks assigned to me. While working for Umbian Inc., I was responsible for producing weekly reports and accommodating internal and external requests. I was also in charge of two web dashboards. In addition, I was able to identify, analyze and propose solutions to a significant problem that I encountered while working at Umbian Inc. This problem was the lack of detailed documentation for the databases the company uses. The proposed solution was to create a web-based search engine for the collected documentation data by using Lucene indexing and search engine [1], which would allow the employees to access and use the documentation system easily and securely.

# Introduction

Being part of the workforce was a great opportunity for me to put all of my academic training to use. Umbian Inc. is a company that provides a very healthy work and educational environment with an open door policy that encourages employees to produce their best work. Consequently, working as a data analyst at Umbian Inc. was a great experience for me, because I had a chance to deal with a large amount of data collected from the sleep apnea patients' devices. Moreover, interacting with employees from different departments also enhanced my social and educational experience. During my time with the company, I was in charge of producing weekly reports for different departments. In addition, I had to accommodate different types of analytical requests from other departments and customers. I also was in charge of developing, updating and maintaining web dashboards. Finally, I identified and proposed a knowledge mapping solution for the company's databases documentations.

# Organization description

## Umbian Inc.

Umbian Inc. is a software deployment company that develops health monitoring solutions. It is owned by ResMed Inc. (www.resmed.com),"one of the global market leaders in the development, manufacturing,

and marketing of medical products for the diagnosis, treatment and management of respiratory disorders, with a focus on sleep-disordered breathing" [2]. The company's flagship is u-sleep CPAP compliance monitoring and management solution. The way u-sleep works is by using the "Management by Exception" strategy, which allows companies to define rulesets for therapy compliance. By doing so allows the system to flag patients who follow or break the rules by trigging pre-set actions and notifications. The system also sends feedback to the patients by using one of the three methods: voice, SMS, and email. These notifications contain automated couching messages for the patient, which helps to improve the patient's therapy compliance. It also provides comprehensive reports to companies regarding the patients they manage. [3]

MyAir is another product that has been developed by the company, which is a patient oriented web application. This web application was designed to give the patient a score out of 100 based on his/her use of the sleep apnea device. The website encourages patients to achieve higher scores that help them achieve better therapy outcomes. This system also provides the patient with feedback and motivational messages based on the device and mask types registered on the website. These notifications are divided into time-based notifications and action-based notification. Time-based notifications rely on research that is done by the company to identify the critical times most patients need couching messages. On the other hand, action-based notifications are triggered by patients' actions, such as praise notifications for complying and sticking to the therapy plans or encouragement messages for patients who are not getting good scores within the system. [4]

## Teams interactions

During the time that I worked at Umbian Inc. I primarily interacted with three teams:

### Customer success team

This team is responsible for checking on companies and making sure that they are utilizing the u-sleep system by setting up patients' correctly. The team does follow-ups based on information from a web dashboard that I was assigned to manage. Making sure that patients are being set up properly ensures that they are getting the appropriate system generated feedback for better therapy outcomes.

### U-sleep team

U-sleep team is the team responsible for developing, maintaining and deploying the u-sleep system for new customers.

### MyAir team

MyAir Team is the team responsible for developing and maintaining the MyAir website.The team also receives an analytical weekly report that summarizes the patients' statistics in the system, which I was responsible for producing.

## Internship work

### Main tasks:

1. Generated and organized weekly analytical reports that went to different departments in the company using SQL and MS. Excel.
2. Accommodated different types of custom requests from customers, Home Medical Equipment (HMEs), and other departments' requests, which involved creating customized and complex SQL queries to ensure the accuracy of the retrieved data.
3. Updated, modified and maintained two web dashboards that display the uploaded data using interactive charts and tables that has been coded using different JavaScript packages such as d3 "Data-Driven Documents", crossfilter and Jquery. Different departments use these dashboards to keep track of sales progress, customers' "HMEs" performance, and system utilization for follow up support when needed.

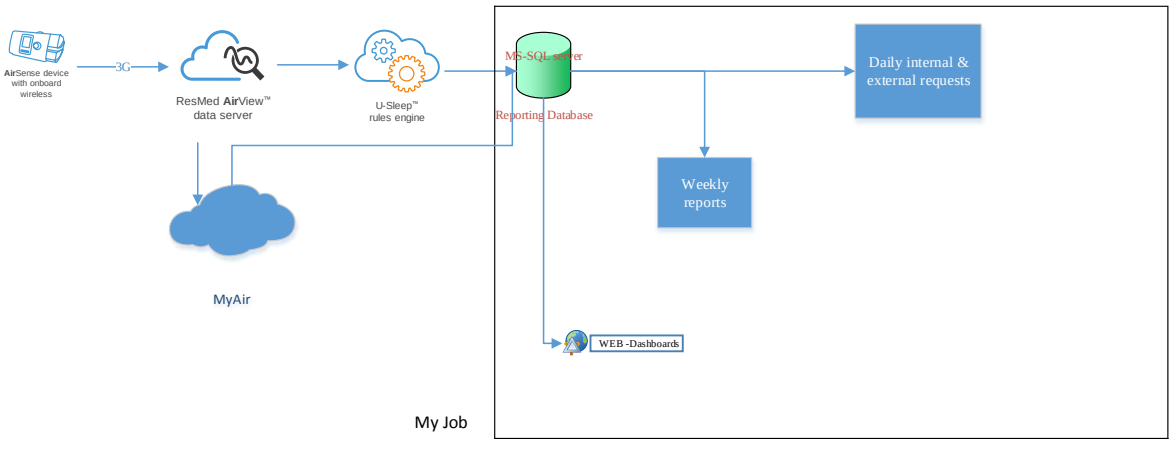


*Figure 1: Internship tasks*

## 1-Weekly reports:

### Sales report:

The sales report is generated by running 3 different SQL queries. After that retrieved data gets formatted in an Excel sheet. Sales reports are mainly used to keep track of the overall sales progress by looking at the number of added or removed patients from the system.

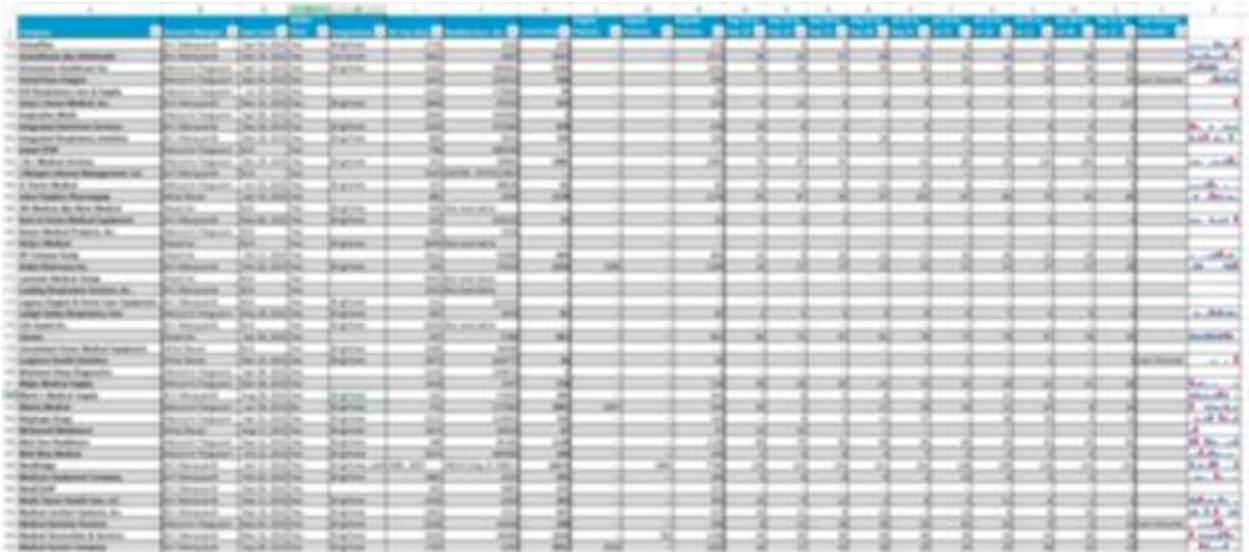

*Figure 2: Sales report (blurred for information privacy)*

### MyAir report:

MyAir analytic report is the most important weekly report because it provides a summary data and visuals about patients' behaviour. Therefore, it is viewed by the top managers at the company. The report can be divided into the following analytic sections:

#### *Compliance and sleep score:*

This section shows the overall compliance numbers and rates the numbers for the past 90 days. It also provides an overall average sleep score for patients since the first day of registration, which gives a good indication of how well patients are using their devices for therapy purposes.

#### *Communication methods:*

This section contains a table with the number of patients who have registered and are using SMS, email, or both as a method of communication. It also provides the numbers for the different statuses for each

method, such as *Disabled* or *Confirmed*. This helps to identify the patients' preferred way of interacting with the system.

### *Devices and masks:*

This section contains some statistics about the devices and masks count used by patients. It helps to identify the popularity of the masks and devices used by patients and it helps the system to send the appropriate feedback and coaching messages to patients.

### *Demographics:*

This section provides a summary of the count of patients' genders registered in the system for statistical purposes.

### *Access method:*

This section provides a summary count of the devices (computer, mobile, or a mix of both) that patients are using to access the system.

### *Alerts and notifications*

This section provides a count of the important notifications and alert messages generated by the system. For example, leak alert, AHI alert, Usage Praise, and badges which help in identifying any red flags regarding the patients' experience with the masks and devices.

### *Marketing statistics*

This section provides simple statistics about the number of patients who gave permission to receive marketing promotions vs. the patients who did not.

### *Patients' website login activities*

This section keeps track of the overall patients' login activities since the day of their registration. This is being done to see if patients who are active on the website are more compliant than patients who are not active.

### *Geographical growth and distribution*

This section provides a summary of the total number of patients being added across the US, Canada, Australia and New Zealand, which gives a clear indication of the patients' growth in each country.

## 2-Custom requests:

During my time at Umbian Inc., I have been assigned 36 requests by my work supervisor, who was responsible for revising the results before sending them back to the concerned parties.

The requests types which the department handled, could be divided in two types:

## Internal requests:

Internal requests can be used for investigating reported problems or exploring new possibilities. These requests could be as simple as asking for patients' count or complex, where different conditions and criteria were involved. For example:

- Finding out the number of patients who change their mask type in the system more than once before July 2015 and the number of patients who did the same thing after that date. This was done to measure the effects of some changes made on the website to ensure that patients are registering the right type of mask in the system.
- Exploring the reasons behind patients opting out from email notification by studying the most recent notification the patient received before opting out.
- Finding out the number of patients who used their devices 90 days in a row since their registration.
- Mask Model: Number and percentage breakdown of the various mask models that users selected upon registration.
- Mask Model Change: The variation (if any) of users changing mask models after registration. Here, the following factors are considered:
    - Number of users that have changed masks.
    - Number of times users changed masks.
    - Time (relative to activation) frame for users changing masks.

## External (HMEs) requests:

When we receive a requests from HMEs, these requests are usually for the purpose of obtaining a summary of all the patients that they have under them, which helps them manage and understand patients' behaviour based on the criteria that they have set.

For example:

- A company requested the following information regarding its patients:
    - Number of days that each patient used the device for more than 4 hours for the past 2 weeks.

- Compliance rate for each patient for the past 90 days.
- Average amount of hours used per day for the past 90 days for each patient.
- Total number of notifications sent to any patient under their care from the past 2 weeks.

In the following format:

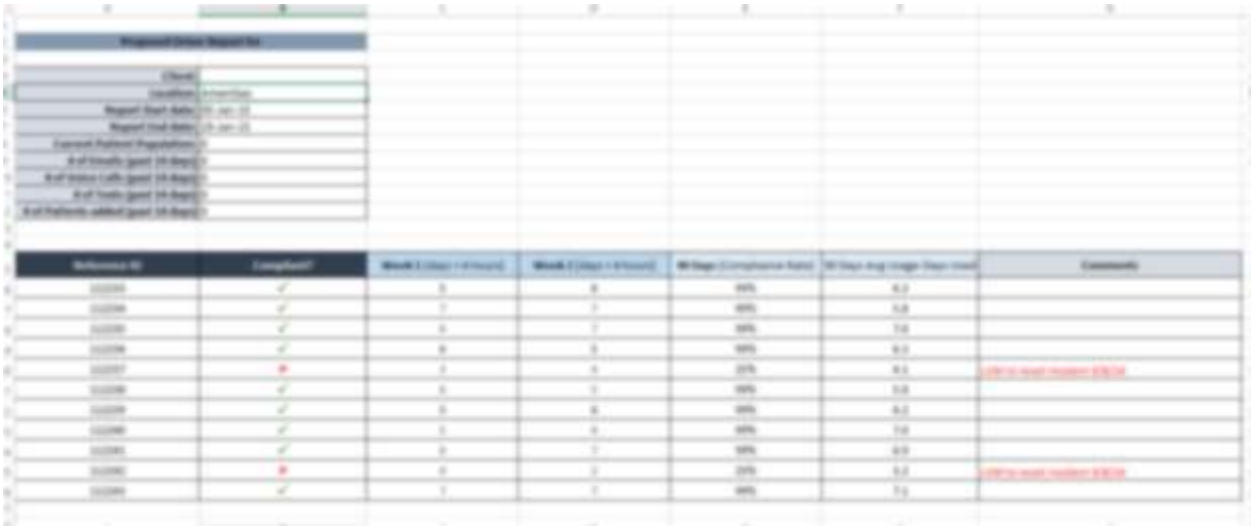

*Figure 3: Sample request (blurred for information privacy)*

## 3-Web dashboard

The main reason of using web dashboards is their ability to deliver up-to-date interactive content efficiently and securely. I was in charge of maintaining, updating and adding new features to two dashboards.

### Sales account dashboard

The sales account dashboard gives the reader summary information about customers, such as company's name, account manager, and the total number of patients. It also provides a graph that shows the number of patients added since the company's registration date in the system.

My main task was to make sure that the dashboard was updated weekly with the recent sales numbers. This was done by retrieving the data from the SQL database and uploading it to the dashboard site. I was also responsible for creating another dashboard that allowed the user to view the company's sales at a specific location with the ability to switch between the locations and company views. I also added a function that allowed the user to export the data tables into an Excel 'CSV' sheet for easy use.

*Figure 4: U-sleep dashboard account report (blurred for information privacy)*

## Customer success dashboard

The customer success dashboard is meant to give the team a summary about the companies' activities using u-sleep system and to makes sure that they are fully utilizing the system. This ensures that the patients are getting the best therapy feedback from the system for better therapy outcomes. This is done by calculating a score for each company based on certain criteria set by the customer's success team. If any company gets a low score in the weekly update, it is contacted by the team for a quick check up and to offer any necessary help.

I was assigned to keep the dashboard up to date by obtaining the appropriate data from the SQL server and uploading it to the dashboard site. I was also assigned to implement multiple changes and updates to the dashboard, which required modifying the existing SQL query to accommodate the new changes. Moreover, I made the necessary changes to the website's html and java script codes in order to display the new features and visuals.


*Figure 5: Customer success dashboard graphs (blurred for information privacy)*

## Relation to health informatics

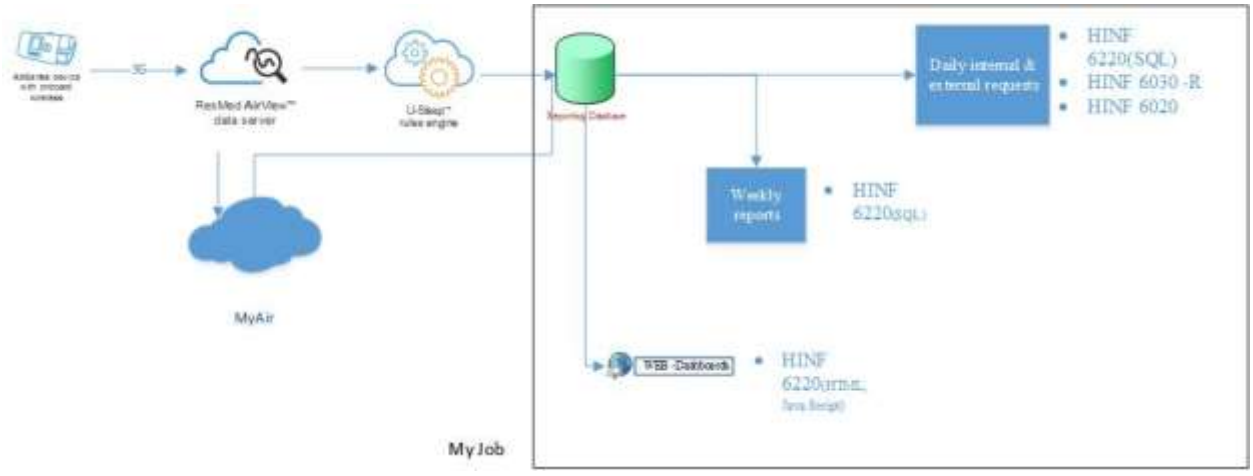


*Figure 6: Academic work in relation to the internship tasks*

Digging through the data being collected from the patients' devices and analysing it to find and identify issues and difficulties with patients' therapy is one of many health informatics desirable outcomes, which was achieved via my role as a data analyst in the company.

- Utilizing my SQL knowledge taught in HINF 6220 - Networks and the Web for Health Informatics course allowed me to create complex SQL statements, which ensured accurate and precise results.
- Using my web development skills and knowledge taught in the same course were very beneficial in creating web interactive web dashboards for an easy and fast use and data visual representation. These dashboards allowed the reader to make the most accurate decisions regarding improving patients' therapy and care.
- Utilizing the statistical knowledge taught in the HINF 6030 class and the research techniques gained from HINF 6020 - Research Methods course helped me to preform analytical and statistical analysis when required.
- The knowledge obtained from the HINF 6110 - Health Information Systems and Issues course allowed me to understand the information flow and the life cycle of the data generated from the sleep apnea devices.

Examples of some requests and how they are affecting patients' therapy:

- Simple match-up between the numbers in the database (of the different types of masks that patients are registering) with the numbers being supplied from the sales departments had identified that patients were not registering the right type of mask on the website. As a result, patients received the wrong feedback messages. This observation triggered changes in the MyAir website design leading to more accurate mask registration.
- Making sure that the HMEs are setting up their patients' accounts properly through the customer success weekly updated dashboard insured the successful delivery of all the system notifications and feedbacks to the patients, resulting in a better therapy outcomes.
- Studying the relation between patient compliance and his/her frequency of using the Myair website to prove that using the website increases compliance (percentage) among patients.

# Problem analysis

During my first week of work, I went through some basic training regarding the company, its products, and more importantly the database structure that I would be retrieving data from. My supervisor did his best teaching and explaining many aspects relating to the company's systems, business, and the teams that I would be interacting with. Furthermore he spent a lot of time teaching me about the databases' different structures and fields. At that time, I noticed that the company lacked detailed documentation relating to the databases. Everyone was relying on their tacit knowledge and no one was recording that knowledge anywhere. Consequently, whenever I needed to find information about a certain field in any of the databases, I had to ask my supervisor. Moreover, if he did not have the information, I had to ask someone in the software development team to get the right information. This experience alerted me to the lack of documentation of the company's large and complex databases.

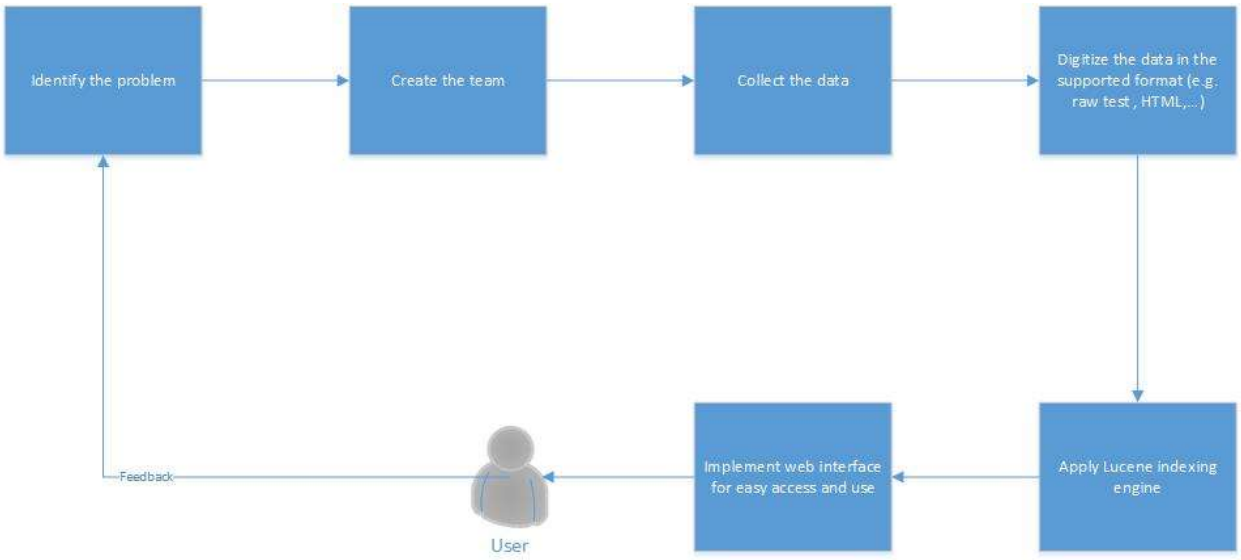


*Figure 7: Knowledge mapping steps*

## The Solutions:

I have proposed a simple knowledge management solution based on the general steps of mapping knowledge [5]. The proposed solution consists of technical and non-technical components. The non-technical part includes all of the managerial and non-technical, related steps. For the technical part I decided to go with Lucene which is "a full-text search library in Java which makes it easy to add search functionality to an application or website. It does so by adding content to a full-text index." [1]. Lucene allows for the indexing of different types of text formats such as raw text files and HTML. Moreover, Lucene adds enables a search of all the indexed data through a web interface.

The following are the steps –involved in the proposed solution:

1. Identify the business problem. In this case, it is the lack of detailed documentation for the existing databases.
2. Get the required approval from the top management by presenting the business benefits that the company will gain by solving this problem. This is a very important step when attempting to implement the policy and when pulling some team members away from their work.
3. Create a knowledge management team that will investigate the problem, as well as develop and implement the solution. This team should be formed by utilizing a member from each team that is being affected by the problem.
4. The team should start the problem solving process by going through the existing documentation and identifying the missing information.
5. Identify the knowledge expert of the databases.
6. Decide the best method for converting the tacit knowledge to explicit knowledge based on the availability of the person that has such knowledge. Data collection could be done via a personal interview, electronic forms, or simply by asking people to hand over any documentations that they have created but never shared with other employees.
7. All the collected data and documentations should be recorded and digitized in Lucene supported formats.
8. Apply Lucene to the digitized files to index them and to make them searchable.
9. Create a web interface for easy search and access using Apache Lucene [6]
10. Asking employees for feedback about the provided service will allow the team to improve the knowledge-based system with time.

## Proper documentation is required for the following reasons:

- Legal liability. Since the company is dealing with different types of medical devices, proper documentation is very important when dealing with legal issues.
- Reduce the training time for new employees, which would lead to an increase in productivity.
- Improve work efficiency and accuracy.
- Reduce informational errors.

## Recommendations

- Create a database change notification process to notify every department working with the same database whenever there is a change in the database fields or structures.
- Hire another data analyst since there is only one person doing this job in the company.
- Prepare a welcome package with all the required reading materials in order to give the new employees sufficient understanding about the company's business and its needs.

## Conclusion

Being part of a very successful health software solution company, like Umbian Inc., gave me a different perspective on the role of health information in the workforce. The experience that I have gained from working as a data analyst was very educational. In addition, working for the company showed me the variety of tasks that I could handle as a health informatician. Being able to utilize many of my technical skills, such as web development and SQL, combined with the statistical knowledge to handle the tasks assigned to me, proved to me that my health informatics skills and knowledge could be applied in the workforce. The most rewarding part of working at Umbian Inc. was the satisfaction of knowing that all the work that I did went towards helping patients.

# References


[1] "LuceneTutorial.com," 2015. [Online]. Available: http://www.lucenetutorial.com/basic-concepts.html. [Accessed 19 09 2015].

[2] "About Umbian," [Online]. Available: http://umbian.com/about/. [Accessed 05 09 2015].

[3] "U-Sleep CPAP Compliance Monitoring and Management Solution," [Online]. Available: http://umbian.com/products/u-sleep-cpap-compliance-monitoring-solution/. [Accessed 05 09 2015].

[4] "myAir," [Online]. Available: https://myair.resmed.com. [Accessed 05 09 2015].

[5] C. O. a. C. Hubert, "Locate your Critical knowldge," in *The New Edge in Knowledge*, New Jersy , John Wiley & Sons Inc., 2011, pp. 108,109.

[6] "Lucene," [Online]. Available: https://lucene.apache.org/. [Accessed 19 09 2015].