Employment Opportunities in HADOOP

Table of Contents

Introduction......................................................................................................................................5

Methodology....................................................................................................................................7

Action Research...................................................................................................................7

Purpose.................................................................................................................................8

Stages of Action Research...................................................................................................8

 Initial Reflection………................................................................................................8

 Planning.........................................................................................................................8

 Action…………………................................................................................................8

 Observation……………................................................................................................9

 Reflection………………..............................................................................................9

Characteristics of Action Research....................................................................................10

Literature Review...........................................................................................................................12

History of Hadoop..............................................................................................................13

Benefits of Hadoop............................................................................................................14

Responsibilities of a Hadoop Developer............................................................................14

Scope of Hadoop................................................................................................................15

Proposal..........................................................................................................................................16

Iteration 1: Research on Hadoop Programming.............................................................................17

 Plan....................................................................................................................................18

 Action.................................................................................................................................18

 Observation........................................................................................................................19 Reflections.........................................................................................................................20

Iteration 2: Select an Institute and Start Training in Hadoop........................................................21

 Plan....................................................................................................................................21

 Action................................................................................................................................22

 Observation........................................................................................................................23

 Reflections.........................................................................................................................24

Iteration 3: Meet with Hadoop Professionals in Better Understanding the Job Landscape and Preparing the Resume……………................................................................................................25

 Plan....................................................................................................................................25

 Action................................................................................................................................26

 Observation.......................................................................................................................27

 Reflections.........................................................................................................................28

Iteration 4: Preparing for interview and applying for the Hadoop developer positions and attending the interview...................................................................................................................29

Plan....................................................................................................................................29

Action.................................................................................................................................30

Observation........................................................................................................................30

Reflections.........................................................................................................................31

Conclusion.....................................................................................................................................32

References......................................................................................................................................33

List of Figures

Figure1: Integration of two organizational schemes for the step-by-step process of action research. Source: (Reason,2008)......................................................................................................9

Employment Opportunities in Hadoop

In this modern era, new technologies are emerging and there are always upgrades occurring. Many of these technologies are used in different fields like robotics, aerospace engineering, and nano technology. Human race has progressed to a greater extent with the use of technologies in various fields.

Data storage and management have gained primary importance recently. Enormous amounts of data are being used and stored by organizations. Storing and managing this huge amount of data has always been challenging a better world even though many related technologies have been rolling out recently. The term “Big data” has emerged and Hadoop technology uses a set of algorithms to process large clusters of data.

This project involves getting a job opportunity as a HADOOP developer so that I can learn many things about the company – what it is, where is comes from, how it can be applied to business processes, and how to get started using it.

With this internship, I would like to do the research job opportunities at Hadoop. The primary purpose for choosing this is for the ongoing latest technology in the software industry. My undergraduate is in computer science and I gained some knowledge on SQL and databases. However, while completing my masters, I have gained good knowledge about big data, and where its types can be uniquely stored and processed in Hadoop. I believe that social media data should definitely need Hadoop for their unlimited competition and real-time decisions that include market share.

Hadoop has its roots at Yahoo!, whose internet web search tool business consists of constant preparing of a lot of Web page information. Eric Baldeschwieler challenged Owen O’Malley, co-founder of Horton works to solve a hard problem: store and process the information on the web in a basic, adaptable and financially plausible way. They took a gander at customary storage approaches; however, they immediately figured out that they simply weren't going to work for the sort of information (a lot of it unstructured) and the sheer amount Yahoo! would need to manage (Baldeschweiler, 2013). It is a data management platform as it offers lower-cost storage framework and an open source development (Yuhanna, 2014). The research suggests that the broad enterprise can embrace and use Hadoop to create big data value (Business Value of Hadoop, 2013).

Hadoop is turned into a foundational engineering at Yahoo and it fundamentals an extensive variety of business-basic applications. Organizations in about every vertical began to embrace Hadoop. By 2010, the group had contracted a large number of clients and a wide undertaking energy had been created (Baldeschweiler, 2013).

While each association is different, their big data are frequently fundamentally the same. Hadoop, as a discriminating bit of emerging modern data architecture, is gathering enormous measures of data over social media activity, click stream data, web logs, financial transactions, features, and machine/sensor data from gear in the field (Business Value of Hadoop, 2013).

According to Baldeschwieler (2013), Hadoop is a system for adaptable and reliable dispersed information stockpiling and transforming. It takes into consideration the handling of vast information sets crosswise over groups of machines utilizing a basic programming model. It is intended to scale up from single servers to a large number of machines, amassing the nearby processing and stockpiling from every server.  **Methodology**

**Action Research**

Action research is a methodology of methodical request that empowers individuals to discover viable solutions to genuine issues experienced in everyday life. Action research has had a long and recognized family that compasses in excess of 50 years over a few continents. Generally, the term action research has been since a long time ago related with the work of Kurt Lewin, who viewed this research methodology as cyclical, element, and shared in nature. Through rehashed cycles of arranging, watching, and reflecting, people and gatherings occupied with activity research can actualize changes needed for social change (Lavery, 2014).

According to Corey (1953), activity research is the methodology by which practitioners endeavor to study their issues experimentally in order to guide, correct, and assess their activities and choices. Understanding action research includes recognizing how the reactions to these struggles helped create new methodologies to and understanding of substantive changes over the long time and across physical, social, and emotional boundaries (Glassman, 2014).

The distinction between action research and other type of research is that during the process, researchers will need to create and utilize a scope of abilities to attain their points, such as careful planning, sharpened perception and listening, assessment, and critical reflection. The traditional research is conducted to report and publish conclusions that can be generalized to larger populations, whereas the action research is conducted to take action and effect a positive change in the environment that was identified. The traditional research can be done in the environment where can be controlled, but, the action research should be done in school and classrooms (Reason, 2008).

**Purpose:** The purpose of the research was twofold. First, it was to analyze how instructors see and translate confirmation accumulated through action research projects led inside the school environments. Secondly, it was to investigate how these understandings could be utilized to educate expert practice inside the schools (Lavery, 2014).

**Stages of Action Research:** The process of action research is just ineffectively portrayed in terms of a mechanical sequence of steps. According to McTaggart (2007), it is generally thought to include a spiral of insightful toward self-reflective cycles of the following:

* Planning a change
* Acting and observing the process and consequences of the change
* Reflecting on these processes and consequences
* Replanning
* Acting and observing again
* Reflecting again, and so on.

***Initial Reflection:*** Action research emerges from an issue, predicament, or uncertainty in the circumstances in which professionals end up. It might be at the level of a general concern, an apparent need, or a course-related issue (Lewin, 1952).

***Planning:*** The most essential outcome of the planning stage is a definite arrangement of the activity that is intended to take or to make the changes (Lewin, 1952).

***Action:*** In the light of experience and feedback, the minor deviations from the plan and record have the deviations with the reasons behind them. Moreover, the new insights can be incorporated in the current project or can be recorded for the future project.

***Observation:*** Detailed observation, checking, and recording empowers us to survey the impact of the action and hence the adequacy of the proposed change. It is better to maintain a dairy or journal to record the observations and insights of the project (Lewin, 1952).

***Reflection:*** At the end of every cycle, it is important to reflect the observations that has made in the dairies or journals.



*Figure 1:* Integration of two organizational schemes for the step-by-step process of action research. Source: (Reason, 2008).

**Characteristics of Action Research:** According to Schuler (1996), the components of action research are of five C’s. They are:

* ***Commitment*** .The factors that all the participants who are involved in the project. The participants need some time to trust each other and to observe the practice, changes, approaches, documents, reflects and finally the results.
* ***Collaboration.*** In action research, the power relations among participants are same; every individual contributes and has a stake. Collaboration is not the same as compromise; however, it includes a cyclical procedure of imparting, giving and taking.
* ***Concern.*** The concern of action research means that all participants will improve a group of critical friends.
* ***Consideration.*** Reflective practice is the careful audit of one's expert actions. Reflection obliges concentration and careful considerations as one looks for examples and connections that will produce significance within the investigation. Reflection is a challenging, focus and discriminating appraisal of one's own conduct as a method for one's craftsmanship.
* ***Change.*** The change is difficult and it is an important element in remaining effective as a human.

In my opinion, action research is a form of investigation designed for the projects to get the employment opportunities as a HADOOP developer. It is a successful tool to get a job in Hadoop. This action research gave me the idea to split the whole process into small iterations. These iterations include the plan, action, observation, and reflection of action research in the particular iteration. This iteration process helped me work efficiently to achieve my targets. It gave some knowledge and skills about the Hadoop and showed me the best way to gain employment.

**Literature Review**

Hadoop is open source reliable and scalable distributed computing platform that stores and process the data. It includes a fault-tolerant storage system known as Hadoop distributed file system (HDFS). HDFS is capable to store large amounts of data, grow incrementally, and survive the failure of major parts of the storage infrastructure without trailing data (What Is Hadoop?, 2008).

Hadoop influences a cluster of hubs to run Mapreduce programs enormously in parallel. A Mapreduce project comprises of two steps: the Map step methodologies information and the Reduce step assembles intermediate results into a last result. Each single cluster node has a neighborhood record framework and nearby CPU on which to run the Mapreduce programs. Information is broken into information pieces, stored over the local records of distinctive hubs, and imitated for unwavering quality. The nearby records constitute the record framework called Hadoop Distributed File System (HDFS). The number of hubs in each one group differs from hundreds to a huge number of machines. Hadoop can likewise consider a certain set of fail-over situations (Lay, 2010).

Hadoop has developed into the system of decision for architects analyzing enormous information in field such as money, promoting, and bioinformatics (Zaharia, n.d.). At the same time, the changing nature of information itself, along with a yearning for speedier criticism, has started interest for new methodologies, counting devices that can convey specially appointed, constant transforming, and the capacity to parse the interconnected information flooding out of social communities and versatile devices (Mone, 2013).

 As Hadoop is a distinct tool, it is aimed at problems that require assessment of all accessible data. For instance, image processing and text analysis usually mandate that every single record be examined, and often infer in the perspective of similar records. Hadoop uses a procedure called Map Reduce to hold out this comprehensive analysis quickly (What Is Hadoop?, 2008)

**History of Hadoop:** In 2002, when project Nutch began as an open source web crawler by Apache Foundation, a working crawler and inquiry framework immediately rose. Doug Cutting, the inventor of Apache Lucene, assessed that a framework supporting a billion file would cost around $30,000. They accepted that it would open up and democratized web crawler calculations. Yet soon they understood that their building design would not scale to billions of pages on the Web (Baldeschweiler, 2013).

Google had additionally confronted the same issue of dealing with billions of website page lists and they made an innovation to defeat this test. In 2004, Google distributed an alternate paper that presented Mapreduce, a parallel programming model focused around utilitarian programming to process conveyed information. In 2005, the Nutch designer likewise made a working Mapreduce usage in Nutch. All the major Nutch calculations had been ported to run utilizing Mapreduce and Nutch distributed file system (NDFS) (Baldeschweiler, 2013).

NDFS and Mapreduce were exceptionally guaranteeing innovations and turn out as an Apache autonomous sub project of Lucene task called Hadoop. In the meantime Doug Cutting offered by Yahoo!, which gave a devoted group and assets to transform Hadoop into a system that ran at web scale. In Jan, 2008, Hadoop graduated to a Top-Level Apache venture, affirming its prosperity. Presently Hadoop is being utilized and upheld by numerous different organizations other than Yahoo, such as Facebook, The New York Times, Cloudera, Hortonworks, and Last.fm (Mone, 2013).

**Benefits of Hadoop:**

1. ***Cost-effective***. Apache Hadoop controls costs by putting away information more reasonably every terabyte than different stages. Rather than thousands too many thousands every terabyte, Hadoop conveys register and capacity for several dollars every terabyte.
2. ***Fault-tolerant.*** Fault tolerance is a standout amongst the most critical points of interest of utilizing Hadoop. Regardless of the possibility that individual hubs encounter high rates of disappointment when running employments on an extensive group, information is duplicated over a bunch so it can be recouped effectively even with circle, hub or rack disappointments.
3. ***Adaptable.*** The adaptable way that information is put away in Apache Hadoop is one of its greatest resources – empowering organizations to produce esteem from information that was already considered excessively extravagant to be put away and handled in conventional databases. With Hadoop, one can utilize different varieties of information, both organized and unstructured, to concentrate more important business experiences from a greater amount of your information.
4. ***Scalable.*** Hadoop is an exceedingly versatile storage platform, due to the fact that it can store and convey substantial information sets across clusters of many reasonable servers working in parallel (Business Value of Hadoop, 2013).

**Responsibilities of a Hadoop Developer:** A Hadoop developer is responsible for the real coding or programming of Hadoop applications. This part is like that of a software developer. The following are some of the responsibilities a Hadoop developer:

* Hadoop advancement and usage
* Stacking from dissimilar information sets
* Preprocessing utilizing Hive and Pig
* Planning, building, introducing, designing and supporting Hadoop
* Interpret complex utilitarian and specialized prerequisites into definite configuration
* Perform investigation of incomprehensible information stores and reveal experiences
* Keep up security and information protection
* Make versatile and elite web administrations for information following
* Overseeing and conveying Hbase
* Test models and regulate handover to operational groups
* Propose best practices/models (Gothai, 2014).

**Scope of Hadoop:** The Hadoop platform has tools that can remove the information from the source frameworks, whether they are log records, machine information or online databases and burden them to Hadoop in record time. It is conceivable to do changes on the fly as well, although more expound handling is better done after the information is stacked into Hadoop. Programming and scripting systems permit complex ETL employments to be sent and executed in a disseminated way. Quick enhancements in intelligent SQL tools settle on Hadoop a perfect decision for a minimal effort data warehouse (Yuhanna, 2014).

**Proposal**

The proposal of my paper is completely dependent on the procedures that I am going to apply to get an employment as Hadoop developer. Finding a job in a good company is not an easy task as it requires a lot of hard work and technical skills to get into the work. This paper provides the action research cycles as it is a step-by-step process in finding a job. A Hadoop developer is a very high-level job, so I have to work hard to get an employment in it. This paper also includes iterations, which is the process to get an employment opportunity in finding a job.

The first iteration requires doing research on Hadoop programming. I would research Hadoop through online websites, journals and books, which gives me the basic idea on Hadoop and its position in the market. While doing this research, I’ll also gather basic information of roles and responsibilities of Hadoop developer. This completes my first step and pushes me little ahead towards my goal.

The second iteration involves selecting an institute and start training. In this iteration, as a part of my action research, I will select one of the best institutes to get trained in Hadoop. My skill set in java is limited and Hadoop is based on java framework. So, I would like to get trained in Java first and then continue to get trained in Hadoop. In this process of searching for an institute, I would like to inquire whether the professor would be able to teach me both Java and Hadoop and also know whether this training would be just theoretical or will it include even practical sessions. If all the above mentioned requirements are covered and I’m satisfied, I would go ahead with the training process. This iteration helps me in having a clear understanding of Hadoop programs and gear up myself for the interviews.

For the third iteration, I would meet with Hadoop professionals to better understand the work culture and some of the roles and responsibilities of a Hadoop developer. As I’m new to the corporate world, it’s difficult for me to analyze the work life of a Hadoop developer. So, meeting such professionals would help me to better understand about the work environment, pay and hours to work. This iteration gives an overall snapshot of the software environment and gives me enough confidence to face the corporate world.

Finally, the fourth iteration is preparing for the interview and applying for the developer position. A crucial step of the preparing is resume preparation. I’ll look for the sample resumes that are available in online and base mine on them. I will start writing my own one. Next to this, I will start preparing for the interview process. For this, I’ll revamp my technical skills that are required for role in the organization. Later, I will go for mock interview questions found online, which helps me analyzing the questioning practice in real time interviews. Apart from technical skills, I would also concentrate on the presentations and discussions as some organizations have included presentations and group discussions as a part of the interview. Last but not the least; I would also like to concentrate on my oral skills as it has a great impact for any role in any organization. This iteration helps me in applying for the Hadoop developer positions through various job portals like indeed.com, monsterjobs.com, and simplyhired.com. Once I receive the confirmation mail or call from the organization to attend the interview, I’ll prepare myself to give the best shot to start my dream career.

**Iteration 1: Research on Hadoop Programming**

**Plan**

After briefly analyzing my proposal on January 30th, 2015, I came to know that smart work and analytical thinking are required to get a job as a Hadoop developer. My initial plan was to research on Hadoop with the help of internet and referring articles. Finding different journal articles, books, and magazines helped to understand the Hadoop developer and its roles and responsibilities. As a part of my research, I examined the Hadoop developer pay scale and its position in the market.

Primarily, the research was scheduled for 3 days and it was an hour session per day. The research of books and journals took about a day and the time taken for internet research and articles was about two hours. I have studied the Hadoop for dummies book to get the basic knowledge about Hadoop. I used [google.com](http://google.com/%22%20%5Ct%20%22_blank) as my search engine. On February 13th, I searched the following:

1. Skills required for Hadoop developer
2. Career growth in Hadoop
3. Average salary for Hadoop
4. Future scope in Hadoop

The search resulted in many web pages, of which I selected just a few. I analyzed and understood these processes to get the best results to know about Hadoop. Later, I filtered the processes and found the best process which met my requirement. Thus, the online research and research on books and journals helps me in understanding the Hadoop developer’s role and also helped me to gain knowledge in Hadoop programming.

**Action**

As a part of my plan, the research took three complete days. It took two complete days to review books and journals just to know the future scope as a Hadoop developer. The articles about Hadoop showed me its current position in market and also the roles and responsibilities of a Hadoop developer. Secondly, I searched [google.com](http://google.com/%22%20%5Ct%20%22_blank) about the Hadoop developer and it directed me to the below link.

<https://www.google.com/search?q=Skills+required+for+hadoop+developer&oq=Skills+required+for+hadoop+developer&aqs=chrome..69i57.16755j0j8&sourceid=chrome&es_sm=93&ie=UTF-8>
the search resulted in 399,000 entries. Since this was too many entities to review, I just filtered these results according to my requirement and analyzed the best processes for my research. The results of my search gave me an idea of what is needed to get trained in Hadoop.

I reviewed these websites and discovered that programming knowledge, self-assessment, and smart work are required for a person to become a Hadoop developer. All these websites suggested learning java to get the programming skills to become a Hadoop developer. These websites showed me the use of Hadoop for “big data” analytics, which is one of the hottest fields in information system today. [Is Hadoop Now Easy to Use?](http://www.quora.com/Is-Hadoop-now-easy-to-use-If-not-what-assistance-do-most-users-need) Clearly explained the need for Support to keep systems up and running for mission critical clusters. I feel that the basic understanding of distributed systems, file system design and java knowledge are required to become a java developer.

Observation

The websites and articles that I have selected were good and helped me by giving useful suggestions in getting a job. The primary and basic observations were the roles and responsibilities of a Hadoop developer. The articles that helped me in getting information about Hadoop are “TOM’S IT Pro real-world business technology”. Among the websites that I have filtered, I felt one website is very useful for me: [Prerequisites for Learning Hadoop.](http://saphanatutorial.com/prerequisites-for-learning-hadoop-2/)

These websites gave me a clear idea of what is Hadoop and the skills required for being a Hadoop developer. These articles and websites have answered all of my questions about Hadoop like:

* How is the career growth in Hadoop?
* What are the prerequisites for Hadoop?
* Is Java and SQL required to learn Hadoop?
* What are the best ways to learn Hadoop?
* From where should I begin training in Hadoop?

From the questions answered by SAP HANA staff, I observed that I have to put some efforts to get trained according to the skills that are required. I know the best way to learn Hadoop after getting the answers to these questions. I have analyzed and understood my strengths and weakness to improve my skills so that I am able to know how good I am in a particular language and how well I have to get trained in clearing the interview.

The third observation was on the managerial requirements for Hadoop that typically enables the basic concepts of Hadoop. Although Hadoop does not require the models, most of the organizations prefer the agile method, as it doesn’t have any specified process.

Finally, I have observed that the technical requirement for Hadoop is java and SQL. It requires both the core and advanced java programming to learn Hadoop. It also requires almost all the concepts SQL like queries, triggers, and cursors.

**Reflection**

After going through the plan, action and observation, this iteration was very successful in achieving my goals and it is completed as per the scheduled time. Performing the actions on time was very important to get a job. Getting a job is not an easy task. Critical thinking, self-assessment, and smart working are required. I have to practice how to speak confidently to face the interview. I believe that I am smart worker and I have critical thinking to solve the problems easily. I feel that in my first iteration, I gave a plan and I was executed according to the plan.

With the previous background of bachelors in computer science, I have basic knowledge on java and SQL. Because of my inexperience, I am worried about getting a job as a Hadoop developer. However, online research helped me discover that there is no experience required for a Hadoop developer. As I am a fast learner, I can able to learn concepts of Hadoop very quickly.

The first iteration concluded with the roles and responsibilities of a Hadoop developer, various managerial skills (like agile method), skills required for a Hadoop developer, career growth in Hadoop, prerequisites for Hadoop developer and also the technical requirements like Java and SQL, which are necessary before getting trained in Hadoop.

**Iteration Two: Select an Institute and Start Training in Hadoop**

**Plan**

After devising a plan on how to gain employment as a Hadoop developer, I came to know that I need some technical knowledge to improve my skills. Initially, I have concentrated on skills like Java and SQL, which are required for a Hadoop developer. As a part of my plan, I have selected a good training institute to get trained. The institute has labs to practice Hadoop. I inquired about the institute and learned about its ranking and its certification before joining the institute. I chose the instructor who had a capability to clear all of my doubts and he is very friendly with the students.

After selecting the institute, the training was scheduled for 60 days with two hours per day. The training was given to a group of people from different educational backgrounds like computer science, electrical, and also from mechanical. Classes were clear and covered basic topics. I attended the classes daily to be on track with the training. In this training I came to learn about the basic concepts of Java, Java programming, Structured Query Language, Big data, Hadoop concepts, and Hadoop programming.

Thus, I selected an institute that provided me an educational environment and logical thinking. Then, the training session began according to the planned schedule. I implemented this plan not only to get the overall knowledge in theories of Hadoop, but also the view of planned action research that helped in understanding the dimensions of Hadoop developer role.

**Action**

As per the plan, I have selected the correct institute with study atmosphere. This institute had many instructors and three instructors teach Hadoop. As I have to get good knowledge of coding, I got trained in Java and SQL first and then I trained in Hadoop.

The training took place for 60 days by Mr. Vivek Meshra, Mr. Srikanth Narayan, and Mr. Ambica Ram. The training started with 10 participants in the class, 5 of which were students with little to no computer background. The first 30 days of training was clearly explained about the basic concepts and also the programming of Java by Mr. Vivek Meshra. I have thoroughly learned the concepts of Java and also the projects of Java. Later in the weekend, my instructor gave me few questions to answer. I have worked to get the logics for those questions by just practicing and reviewing the notes given by instructor. After getting trained in Java, I wrote a Java certification and I scored well in it.

Secondly, the SQL training took 10 days by Mr. Srikanth Narayan. In first week, I learned about the concepts of SQL, which included the structure and commands (DDL, DML, and DCL). He used to give daily assignments after discussing the queries in the class. As I am from computer science background, it was a little easy for me to write a query. Later in the second week, I learned about the SQL server and normalization concepts.

 Finally, the last 20 days of training was given by Mr. Ambica Ram. In the first week, I learned the basics of Hadoop like its importance, the main components (HDFS and Mapreduce), and its overview. In the second week, I learned the working of Hadoop Distributed File System (HDFS). In the third week, I learned the abstraction of Mapreduce which included word reduce code, its failures, and recoveries. In the fourth week, the instructor explained the Java Map method, Pig Latin scripts and basics of HiveQL. Thus, after getting trained in an institute, I gained knowledge not in Hadoop, but also in java and SQL.

**Observation**

After the second iteration - the training on the concepts of Java, SQL, Hadoop and the role of a Hadoop developer - there were many observations that came to my mind. As the plan dealt with two agendas, the second one took the major contribution. Moreover, the observations were made for better understanding the concepts of Hadoop.

 The first basic observation was about selecting the institute. This observation was made for training in Hadoop. I observed that the institute should have a lab facility with all the software installed on the computers. I also observed that the institute should provide the Java, SQL and Hadoop training with the respective instructors to become flexible with the Hadoop.

The observation I have made with the instructors is the instructor should be Java certified and should provide the material for Hadoop.

 The second observation was on technical skills of Hadoop. Apart from facilitating the departments, a Hadoop developer mainly deals with writing the reliable and managerial coding, good knowledge of database structures, and back-end programming. Thus, a Hadoop developer should be ready to write Hadoop as well as Java and SQL coding.

The third observation was about training in Hadoop. The training was done in three parts:

* ***Java training***: The observation I have made from this was to practice the programs and understand the object oriented concepts in Java.
* ***SQL training***: I observed that the theories, practices, and principles plays a major role in getting knowledge of SQL.
* ***Hadoop training:*** The observation I have made from Hadoop classes was that the Mapreduce, HDFS, and Pig Latin script played a major role in getting skills.

 Finally, the observation I made from this iteration was that gaining the knowledge in each language was an added advantage in the software industry. A Hadoop developer is not expected to master in SQL and Java but a little knowledge of coding is very important.

Reflection

The second iteration was successful with the goals and also with the scheduled training. On selecting the institute, on-time training schedules, interactive sessions and assignments, this iteration has added one extra feature than the expected outcome. However, selecting an institute helped me in knowing the institute provides some of the Java certifications which were an advantage for me to upload on my resume.

As I have a computer science background, training in Java and SQL was not difficult for me. So, I was able to complete these courses quickly. During my bachelor’s degree, I had core Java and SQL courses. Thus, I trained in Advanced Java, which was very new to me, and also more principles of SQL that were helpful.

As per the planned proposal, the training was expected to be more practical on Hadoop. The Java and SQL seemed to be practical but Hadoop was not that practical as expected. This training opened up a necessity for new tools, which sometimes are required for the Hadoop developer.

However, the second iteration was successful and ended up with an overall understanding of Hadoop, technical requirements like Java, SQL and Hadoop and also some of the certifications of Java, knowledge on some of the tools like Pig, hive that has a flexibility to understand on hierarchy levels.

**Iteration Three: Meet with Hadoop Professionals in Better Understanding the Job Landscape and Preparing the Resume**

**Plan**

After getting the technical knowledge, the next step was to meet the Hadoop professionals to understand the scenario going in the software industry. Finding the Hadoop professionals on social networking sites like Facebook, LinkedIn gave me an idea of their roles in their work. As LinkedIn is a professional networking site, I was able to find many software professionals with this site. Sharing their experiences helped me to improve my skills as well as body language. By contacting them, I could know the growth in income after some experience. Meeting with Hadoop professionals gave the better understanding of job market and could be able to know the responsibilities as well as how to do a Hadoop project.

 As a part of the plan, the employer should know my technical skills to get a job. So, I prepared a resume to apply for the interviews. Preparing a resume was not an easy task and it took around four days to complete. The resume was written to effectively communicate my assets to an employer. Mentioning the accomplishments in the resume gives the employers an idea of my academic success. As resume is a self-marketing tool, it is presented in a clear and easy format to apply for a job. I prepared a resume according to the company requirements based on my skills that included both technical and managerial skills. The resume was written by performing the tasks like surfing the internet to get sample resumes, including the skills, and experience in resume, checking for grammatical mistakes, and reviewing the resume before applying.

Thus, I met with the software professionals with Hadoop experience and learned the managerial skills from their experience. I implemented this plan to get the clear understanding of work that is done by Hadoop developer and also for preparing the resume. The resume was analyzed, prepared, and reviewed in the action part of the action research.

**Action**

According to the plan, I met with the software employees through Facebook and LinkedIn and requested them to share their experiences. The Hadoop professionals gave me a clear idea of salary structure and their responsibilities while doing a project. Because of my inexperience, I had no idea of how to communicate with the colleagues, team leaders and managers in the office. By talking to Hadoop professionals, I have a better understanding of how to communication between colleagues and superiors.

 Secondly, the resume was prepared in four days as per the planned schedule. In the first day, I browsed the internet to get the sample resumes to apply for a job. I learned about the format of a resume and listed all of my strengths and weakness roughly. The strengths and weakness included technical knowledge and managerial skills.

The second day focused on writing a resume by including the career objectives, strengths, and weakness as per the position of job. The career objective was a powerful opening for a resume. It included the actual goal in the particular organization and the position available in it. The objective was two or three lines in length.

 On the third day, I presented technical and communication skills according to the company requirements. I also presented the accomplishments in the resume. I have six months experience on working in mini project, which was mentioned in the resume. The mini project was Optimizing Wireless LAN for Long wall Coal Mine Automation, which was a Hadoop project. The project title, description, duration, and software were mentioned in the project for better understanding my experience as a Hadoop developer.

Finally, I checked the grammar mistakes and reviewed the resume before submitting. The mistakes were identified and fixed to make a perfect resume. Even though the mistakes were fixed, it was better to show the resume to an experienced person in this field. Therefore, I forwarded my resume via GMAIL to Mr. Ram Movva, an experienced Hadoop developer. He checked the resume and fixed the mistakes and sent it back to me. Now, the resume was good for marketing.

**Observation**

After implementing the plan, meeting with Hadoop professionals and preparing the resume, there were many observations made. From this iteration, preparing a resume took a lot of time as it is very important to get a job in any field.

The first observation was about my way of thinking about the Hadoop developer’s role. There were a lot of differences in my way of thinking after meeting the Hadoop professionals. Previously I thought that a Hadoop project was same as a Java project and there was no increase in the income. Now, I learned that Hadoop was quite different from the Java project and the pay was high for Hadoop employees. The second observation was communication in the office. By contacting the Hadoop professionals, I observed that there was a lot of communication between colleagues to better understand the project.

The third observation was on the format of resume. By looking at the sample resumes through internet, I observed the contents that needed to be included in the resume. The fourth observation was on the career objectives and the weaknesses. After having a look at the sample resume, I started analyzing my strengths, weakness, and career objectives. I observed that the career objectives in my resume were the standout for the competition and should be adjusted for each job application to match the position. The fifth observation was about checking the mistakes written in resume. After writing the resume, I reviewed and fixed the mistakes. I thought it was a perfect resume, but after forwarding it to an experienced person, I learned that there were grammar errors that I had missed previously.

Finally, I observed that preparing a resume was like a mirror which can show my skills, capabilities, and experience to the employer. So, I prepared my resume strongly with no errors and forwarded it to experienced person to fix the errors.

**Reflection**

This iteration was successful by meeting the experienced persons in Hadoop and preparing a perfect resume. Contacting the Hadoop professionals, sample resumes, and fixing the errors helped me to realize the importance of showing my skills to the employer. Because of my inexperience, I have knowledge of the office environment. After talking to the experienced people, I came to know many things related to Hadoop, most important of which was communication skills.

As I had six months experience in the mini project, I clearly understood the team work in the project. The prospective employer knows the relevant experience in Hadoop and how I qualify for the job by seeing the resume. I realized the importance of project and purpose of resume. The perfecting of resume gave me confidence as it plays well in the market. By providing the accomplishments in resume, I was able to provide the generic description of myself. I realized the important of accomplishments and those were highlighted in the resume. These accomplishments of doing a mini project were clearly presented.

Reviewing the resume was just as important as forwarding the resume to an experienced person.

**Iteration 4: Preparing for Interview and Apply for Hadoop Developer Position and Attending the Interview**

**Plan**

 After preparing the resume, the next step was to prepare for the interview and apply for a Hadoop developer position. So, I have to prepare well for the interview. This interview is a platform for me to give the interviewer an idea of how can I work, and show the skills and knowledge I have about Hadoop.

 Initially, the first plan that I have was to identify the most common questions asked during the interview. The second plan was to focus on my strengths and to brush up on the technical knowledge for which I got trained. The third plan is to improve the communication skills by managing the interview tension and excitement. The fourth plan is to ask seniors or experienced employers to conduct mock with me. The fifth plan is to review the mistakes I made in this process and fix them.

The next plan was to research possible Hadoop developer positions through the websites like monster.com, careerjobs.com, and indeed.com. The last plan is to apply for Hadoop position that is available in the job searching websites. If my resume is selected for an interview, then the company may give me an aptitude test. Some companies do not have any aptitude test. If I have cleared the aptitude test, I have to attend the interview and the interview might be face-to-face or telephonic.

Finally, I have prepared for the interview which includes questions about problem solving and technical skills. Then, I apply for the Hadoop developer position through websites and attend the interview with confidence.

**Action**

As per the plan, I prepared for the interview. Using the internet, I searched for the most common interview questions. I reviewed the technical skills that I obtained while training during the 2nd iteration. The technical skills included the skills that I included in the resume. I prepared for the communication skills by standing in front of the mirror and reciting a presentation on Hadoop. I also gave the explanation of Hadoop programs in front of my roommates to improve the communication; this also helped build my confidence.

Then, I looked for jobs in Hadoop developer positions through internet. Then I shortlisted the companies based on vacancies, company reputation, location, and position. Then I uploaded my resume and my personal details in the company website.

Finally, I got a call for an interview. The interview was face-to-face and it was located in a location which is very far from me. While going to an interview I was confident enough to answer the questions and the interviewer was impressed with the answers. Later, I got a call from the employer and informed me that I got a job as a Hadoop developer.

**Observation**

After implementing the plan, preparing for interview, and applying for a job, I observed that I am confident enough to present the practical skills but I am not strong at academic skills.

The observations I made were about the qualities required to impress the interviewer. Previously, I have no idea of the qualities but after preparing for the interview, I became more confident. I observed that the competition was required to crack the interview. I noted how far better qualified I was above the other candidates. The mock interviews really helped with this.

The observation I made about applying for a job was the company’s profile and its requirement, which match my skills. I observed how I can build my career after joining the company. After getting shortlisted for the interview, I observed that concentration is required while attending the interview.

**Reflection**

This iteration was successful in preparing for the interview and cracking the interview in a big company. It helped me in patching up the strengths and weaknesses. The mock interviews and suggestions given by the seniors helped prepare me for the interview. I learned to convert all of my negatives to positives.

As the selection of a company has a crucial influence on my future and the career growth, I was very careful in selection. I found very good job vacancies and took additional care in matching the company requirements and my skills. As I have concentrated pretty much on practical skills, I was very confident in getting a call for the interview. I attended an interview and was very confident in answering the questions. Finally, I accepted a Hadoop developer position in a company.

 **Conclusion**

 This action research project is successful with a knowledge gaining experience. In these 15 weeks of my research, I have gained knowledge of getting a job. It made me believe that smart working is necessary to crack the interview. In my project, I have learned about the process of finding a job, its methodology, and action research. With the help of action research, I realized the process to adopt, how to plan, and what should I plan. It also made me to learn how to divide the framework into small segments and work collectively with those segments.

 In this paper, I have divided the action research into four iterations and started working on these iterations. These iterations helped me search for the employment, work on mini projects, meet the Hadoop professionals, prepare the resume and interview, apply for the company, and attend the interview. These iterations helped me to split the work to achieve my goals. Thus, I prepared a good resume with the help of Hadoop professionals and prepared for the interview. Finally, I submitted my resume in the market, attended the interview, and got a job as a Hadoop developer.

 References

*Business Value of Hadoop.* (2013, June). Retrieved from Hortonworks Inc.: http://hortonworks.com/wpcontent/uploads/2014/05/Hortonworks.BusinessValueofHadoop.v1.0.pdf

Corey, S. (1953). Action research. In *Action research to improve school applications.* New York: Bureau of publications.

Dumbill, E. (2012, January 19). *Volume, Velocity, Variety: What You Need to Know About Big Data*. Retrieved from Forbes: http://www.forbes.com/sites/oreillymedia/2012/01/19/volume-velocity-variety-what-you-need-to-know-about-big-data/

Baldeschwieler, E. O. O. (2013, may). *Apache hadoop basics*. Retrieved from Horton works: http://hortonworks.com/wp-content/uploads/downloads/2013/07/Hortonworks.ApacheHadoopBasics.v1.0.pdf

Freeman, W. (2002, Febraury). *Business Resumption Planning: A Progressive Approach*. Retrieved from SANS Institute: http://www.sans.org/reading-room/whitepapers/recovery/business-resumption-planning-progressive-approach-562

Glassman, M. (2014). Participatory action research and its meanings: Vivencia, praxis, conscientization. *Adult Education Quarterly*, 206-221.

Gothai, E. B. P (2014). A Novel approach for partitioning in Hadoop. *Journal of Theoretical and Applied Information Technology*, 537-541.

Lavery, G. S. (2014). Action research: Informing professional practice. *Issues in Educational Research*, 162-173.

Lay, P. (2010, november). *Leveraging Massively Parallel Processing in an oracle environment for big data analytics*. Retrieved from An oracle white paper: http://www.oracle.com/technetwork/database/bi-datawarehousing/twp-hadoop-oracle-194542.pdf

Lewin. (1952). Stages of an Action Research Project. 462-463.

McTaggart, S. K. (2007). Participatory action research. *Communicative action and public sphere*, 271-327.

Mone, G. (2013). Beyond Hadoop. *Communications of the ACM*, 22-24.

Reason, P. A. (2008). Participative inquiry and practice. In *The SAGE Handbook of Action Research.* london: SAGE.

Rosenthal, D. P. (n.d.). *Business Resumption Planning*. Retrieved from The business forum: http://www.bizforum.org/whitepapers/calstatela.htm

Schuler, B. E. (1996). Action research in early childhood education. 3.

*What Is Hadoop?* (2008, March). Retrieved from cloudera: http://www.hurence.com/sites/default/files/What%20is%20Hadoop.pdf

Yuhanna, M. G. (2014). The Forrester Wave™: Big Data Hadoop Solutions, Q1 2014. *For Application Development & Delivery Professionals*, 15.

Zaharia, M. (n.d.). *Introduction to Mapreduce and Hadoop*. Retrieved from RAD Lab: http://www.cs.berkeley.edu/~matei/talks/2010/amp\_mapreduce.pdf