Vaacha (Voice): A Tribal Healthcare Management System

Ranjit Singh
Dhirubhai Ambani Institute of Information and Communication Technology,
Gandhinagar, Gujarat, India

Report on the final year BTech Project
Research Supervisor: Prof. Binita Desai
May, 2007

Vaacha – a Tribal Health Care System

Ranjit Singh
Research Supervisor: Prof. Binita Desai

Abstract – **Healthcare in the 21st century is not a reference index for the support system for patients suffering from various diseases occurring in particular regions. It serves the purpose of evaluating quality of life in a society. In under-developed regions, the impact of the failing health of an individual isn't just a function of his abilities to support his family; it also is a function of the family's ability to support his healthcare expenditure. It reflects upon the ability of the family to sustain itself with the loss of one working member. Thus, the tribal imagination requires not just an affordable cure for diseases but also a methodology of preventing outbreaks of diseases wherever possible. Building a Knowledge-based HealthCare Organization around the patients that come to the Health Camps organized by Bhasha, an NGO working for Tribal Development in and around Tejgadh, is a step towards understanding the circumstances which lead to generic disease outbreaks in the Tribal Belt of Gujarat and creating visualizations by which one may evaluate situations and strategize prevention methodologies.**

Index Terms – Visual Design, Health Care Organizations, Sustainable Healthcare Services, Geographical Information Systems.

1. Introduction

Healthcare in the 21st century is facing three very large forces of change; namely, an informed and empowered consumer, the need for e-health adaptability and a shift from focusing on primarily curing diseases to the prevention of diseases. But these changes are stereotypical to cities. The definition of healthcare facilities alters as one moves towards the lesser developed regions. The informed and empowered consumer is replaced by individuals who lack information and require sustainable healthcare support facilities. The need for e-health adaptability is replaced by the need for health camps at places that are closer to villages and hence are more accessible. The shift from focusing on primarily curing diseases to the prevention of diseases is the only change that remains constant.

Prevention irrespective of the economics of an area is always a function of the amount of knowledge that one has about the area. Knowledge about the area is again directly related to the amount of information there is about the people of that area. The project has been initiated with the assumption that a proper mix of societal, occupational, cultural, economic, and health information would create a system that can be actively used for a holistic idea of the dynamics of a region and serve the purpose of effectively evaluating policy matters with respect to Tribal life in and around Tejgadh.

The rest of the report is organized as follows: Section 2 provides the problem definition. Section 3, provides the outline of the persona of various probable end-users. Section 4 deals with the conceptualization of the System Design. The division of the solution into categories and the methodologies used for solving each of these is dealt with in Section 5. Section 6 provides an evaluation of the expected impact of the solution. The report concludes with an outline of the future work required for integrating the methodologies discussed in Section 4 and creating the complete solution in Section 6.

2. Problem Statement

Bhasha, an NGO working for Tribal Development in and around Tejgadh has been organizing regular health camps primarily in the Vadodara district since 2002. Along with collecting data in relation to patient health, they have also been making surveys regarding land records, occupational patterns, migration patterns, malnutrition and surveys about specific diseases such as Sickle Cell Anemia, Silicosis, Tuberculosis, Pellagra and Handicaps. It has its own set of constraints as to the amount of investment it can make into the project and it has a set of dedicated field workers who are ready to respond to change and train themselves for any new technological solutions.

Given the diversity of the kind of surveys being done by the field workers in the region, the problem at the first level is of effective digitization of the survey data. The data being collected is dynamic in nature. It changes with every survey. Even the surveys go under constant improvisation and alterations. The only constant that can be worked with is that the data is always of the people in the area of operation of the NGO. Thus, the forms being created require provisions of constant changes in the way data is being updated into the system, alterations in the table columns and support for creating new variables about which the data needs to be collected.

Once the data is digitized, the second level of the problem is about identifying various methodologies of viewing the content being uploaded into the system. The approach might be in terms of *views* created in the database which are more statistical and could be used for data mining. Or it could be looked upon as a problem in information visualization wherein new methodologies could be used to view data in terms of Family Trees, Graphs or any novel approach to data visualization.

Among many such methodologies of information visualization, one might also look at development of a GIS as a probable way of looking at the data at end, collating it and representing it in terms of land boundaries. Matching the relational data being collected with the GIS data requires a mapping platform such as *Oracle's Spatial extension*[1]. Here one should note that solutions have been devised to compensate for the lack of a mapping platform given that the Latitude and Longitude of the regions being referred to are available. Thus apart from creating secondary statistical representations, building a GIS for Healthcare forms the third level of the problem statement.

The solution here should be looked upon as a potential web-service with update, data-entry and critical views of data being reserved for select end-users. While the generic representations and information is open and accessible. The GIS could be used potentially to evaluate problem areas and devise solutions with region-specific policy making. It could be used to create awareness within the Tribal community and by the Government for policy-making in Tribal regions.

3. End User Persona

The solution requires an evaluation of the criticality of the information that it can contains about the Tribal Life in general and how it should be used for the general welfare of the people it is trying to serve. The question regarding the generic accessibility of this information mostly revolves around the ethics of its possible usage. The end-user persona presented below is given by the client and the solution will provide limited access to information for general users as per the request of the clients. The users that are targeted by the solution are as follows:

- **Field Workers at Bhasha**: Field Workers are responsible generating data, conducting surveys, updating information as well as surveys whenever necessary. Their proficiency in English is a matter of concern because usually the field workers are proficient in Gujarati and the surveys being conducted are also in Gujarati. Thus the solution might need to be multi-lingual at least at the data entry level. Beyond data entry, they are also the people who have hands-on understanding of the geographic locations being mapped by the solution and would provide the necessary feedback to ensure that the GIS and the other representations are working as accurately as possible. Thus the field workers have a multi-point access into the system and are the primary people that would ensure that the data being collected and displayed is as precise as possible.
- **Doctors**: Healthcare requires medical history of patients. The doctors at the health camp refer patients with serious illness to hospitals in the city of Vadodara and other Public Healthcare Centers. The system aims to provide access to a specific set of doctors being referred to, so that they can access patient histories and check for initial food habits, symptoms, diagnosis and treatment and also update such information. The access is limited to Medical Report of patients.
- **Policy-Making Authorities**: Policy-making requires a holistic understanding of a region in terms of its economics, health, culture, geography and development patterns. The data being collected is expected to provide an evaluation of the quality of life of Tribal People in general and the representations being provided are expected to serve as ways to look at such data and make conscious decisions about Tribal welfare. The system is expected to provide complete access to representations being created on the data to the Policy-Making Authorities. Thus, the users in this domain can only access information to view it and would not be able to update or insert any new data into the system.
- **General Usage**: The solution will provide for restricted access of information to general users as and when decided by the Client.

4. System Design

The process of devising a solution for the given problem statement started with an initial focus on developing concepts that would be used throughout the implementation process. The solution was devised as a people-oriented system which would look at the Tribal People as the main focus object. Hence, the rest of the concepts were developed around people.

The index of quality of life of people is dependent on the set of living conditions of the Tribal people. Hence with people in the center, their economic condition, health, education, society, culture and occupation form the outer circle of concepts around which the rest of the connections with other concepts are developed.

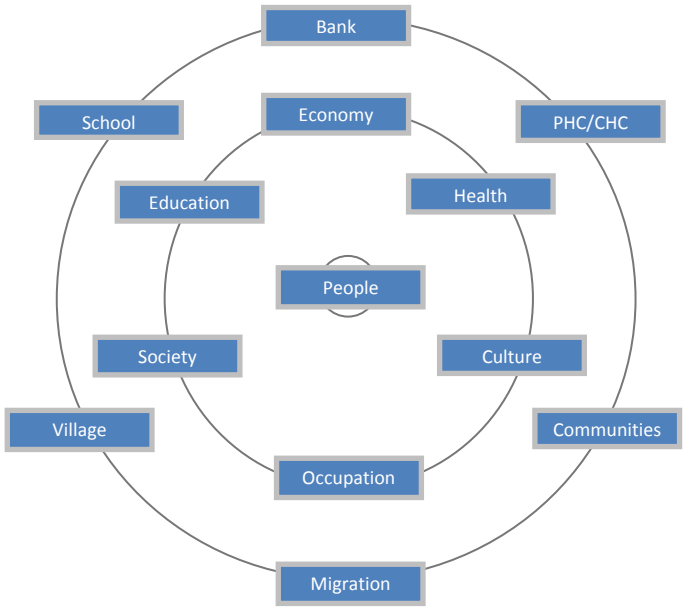

FIGURE 1
RELATIONSHIP DIAGRAM BETWEEN CONCEPTS FOR SYSTEM DESIGN

Hence, as shown in figure 1, the system incorporates the idea that people have an economic status which connects them

to a Bank or they have a society which in turn connects them to the village they live in or they have a set of occupations for which they need to migrate from one place to another.

In the process of creating these concepts and relationships between them, a *concept map*[2] for the solution was developed. A concept map is a diagram for visualizing the relationships between different concepts. Every major concept in the diagram would have a related set of properties and connections.

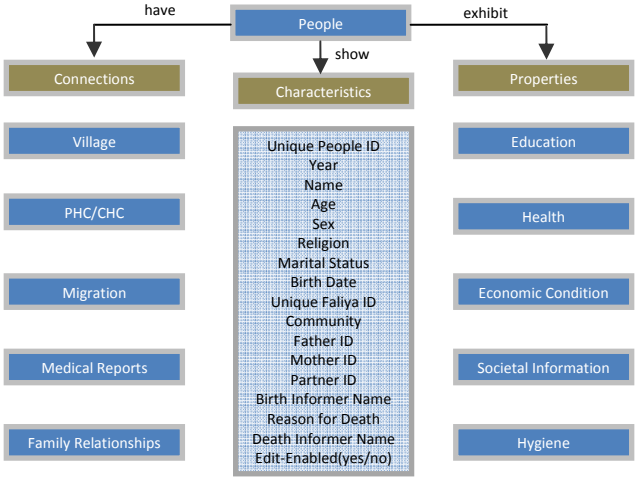

FIGURE 2
PART OF A CONCEPT MAP SHOWING RELATIONHIPS BETWEEN PEOPLE AS A CONCEPT WITH OTHER CONCEPTS (MARKED IN BLUE)

In terms of *connections*, people are connected to their village, public health center, medical reports, family relationships, migration, school etc. and in terms for their *properties* they are connected to education, health and various other indices of quality of life. The concept map also incorporates the notion of *characteristics*, which incorporates the features that constitute an integral part of a concept. For example as seen in figure 2, people as a concept have characteristics like name, birth date, parents, wife etc.

The advantage of creating such concept maps is that they directly correspond to the schema of database management system being put into place. The concept such as *People* correspond to table within database with name *People*, the characteristics correspond to the column names within the table *People*, the properties of the concept *People* corresponds to a set of tables which have the same primary key as that of the table *People* and connections of the concept *People* corresponds to tables that can be referenced using the primary key of the table *People* as a foreign key.

5. APPLIED METHODOLOGIES

The solution presented in this report is expected to be used as a fully functional prototype of the final solution being envisioned. The solution has a three-level development model. First, taking into account the dynamics of the data one needs to create a data-entry mechanism for a database management system based on the surveys being conducted. Second, requires working with presentational solutions such as *Flash* to create visualizations of the collected data. In effect, one has to create ways of looking at the data. The last part involves mapping this data to the GIS data to create representations based on land boundaries matching people to land where they live.

I. Data Entry

The initial problem with data-entry starts with whether the solution being created is multilingual or not. Support for Gujarati Unicode should make *Oracle* as the possible choice for creating the database. This in turn would involve expenditure in terms of licensing whereas training the Field Workers for data-entry in English would make it possible for the system to work entirely in English and hence *MySQL* can be used instead of *Oracle*. The database has been developed in *Oracle* with support of multilingual data entry mechanisms.

Presently the data being worked upon is basically derived from the Primary Census Abstract of the CENSUS conducted in 2001. Along with this data, the excel sheets created by the field workers of Bhasha has also been taken into account. The rest of the CENSUS data concerning migration, occupation, culture and household amenities has also been looked into.

The data being collected is people-centric. Hence, the *primary key* for most of the tables is going to be the *Unique People ID* that is automatically generated for each entry into the People Table. Mostly, it would be a function of an individual's serial ID combined with his state ID, district ID, taluka ID, village ID and faliya ID but since we do not expect people to remember a 15 digit code, hence people would be identified based on their name, village and age information. Since this information might be same for two individuals the users would be prompted to verify which person is being referred to and then the records might be updated, created or manipulated. The schema of the database has been reviewed and approved by the Client side.

To support the dynamic nature of the data and surveys being entered, updated and worked with, usage of static entry forms would serve limited purpose because they might end up dysfunctional after secondary reviews by the NGO. Thus, the system created has provisions of providing fully-flexible data entry methods while internalizing the *Primary Key* generation so that the uniqueness of the data might be preserved. Thus, ways of creating survey forms, putting data into various tables from a single survey entry, supporting addition of new columns, tables and surveys and finally enabling users to access data in survey formats have been devised.

This has been accomplished by creation of an additional description table for each table in the schema and a survey forms table which consists of every unique column in every table of the schema. Using these tables, survey forms can be generated and used for data-entry.

To cite an example as in Figure 3, the *survey_forms* table containing all columns of the *people, village and phc* table accesses *describe_people, describe_village and describe_phc* to create a UI to display the kind of information required of

the user. Once the forms are created using the information in the *survey_forms* regarding the columns into which the data is to be entered and their respective description in *describe_people, describe_village and describe_phc*. The data entered by the user is entered into the *people, village and phc* table. Thus, one can choose which columns to enter data into in a table and also which tables to choose the columns from if the *survey_forms* contains unique columns of different tables.

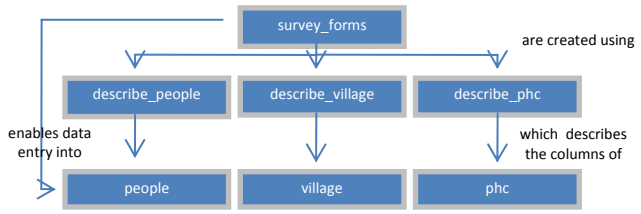

FIGURE 3
REPRESENTATION OF RELATIONSHIP BETWEEN THE ADDITIONAL TABLES ADDED INTO THE DATABASE SCHEMA

II. Creating Views for Collected Data

Data that was initially digitized by the Field Workers themselves was mostly in *Excel sheets*. The formats were specified for data entry and patient histories were documented in terms of a few set parameters. The problem encountered in the method was that creation of graphs and charts wasn't feasible because of the choice of parameters and hence the ways of looking at data was purely statistical column wise entries into the file.

A shift to a database management system and creating *views* for various evaluation parameters enabled a better interpretation of the data. For example, comparing the income levels of people suffering from Sickle Cell Anemia would create a better understanding of the quality of life of patients and evaluating strategies to improve their standard of life.

The visualization mechanism may involve more than just number crunching for statistical evaluation. To evaluate patterns of marriages to reduce the spread of a genetic disease like Sickle Cell Anemia might involve representations in terms of Family Trees. If one can define semantically the scale of problems, one may also create a representation for areas that are affected by a particular problem based on the gradient of the scale.

Taking the idea of the scale of problems into account, one has a listing of various attributes to the level of villages in the primary census abstract of CENSUS 2001. The scale is defined in terms of ratio with the total population of an area. For example, there are attributes such literacy percentage, working population percentage, cultivators among the working population percentage etc. Since during comparisons one may like to look at a variety of attributes and choose between them and there should also be an option for choosing the number of villages being compared. The problem was to figure out ways of giving that amount of flexibility to the user while comparing the available data.

The solution to the problem can be visualized with the origin being treated as a circumcenter. Given that there is just a single attribute to be compared; the axis of the attribute could be any line that defines the radius of the circle. If there are two attributes to be compared; the axes of the attributes would define the diameter of the circle. Given that there are three or more attributes, the axes would just be lines connecting the vertices of an n-sided polygon to the circumcenter/origin. The villages that are being compared would be placed on these axes with their distance from the origin being in the same ratio as the percentage value of the attribute. For example, a village with 50% literacy rate would be placed mid-way between the circumcenter and the vertex on the axis that stands for literacy rate.

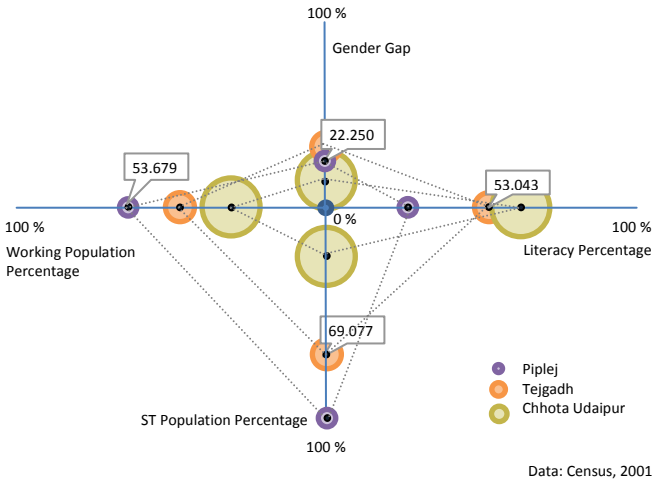

FIGURE 4
VISUAL REPRESENTATION OF COMPARISON BETWEEN THREE VILLAGES IN TERMS OF THE ATTRIBUTES RELATING TO LITERACY, GENDER GAP, WORKING POPULATION AND ST POPULATION

As seen in Figure 4, Piplej, Tejgadh and Chhota Udaipur are being compared in terms of four attributes and the dashed lines connect their respective places on every axis. The radiuses of the circles that represent these villages are in correspondence with the population of the area. The net result is that we obtain a polygon that represents the state of every village in terms of the attributes that are selected. The state of every village when compared relatively to each other could provide the necessary insight for policy making. As an example, Chhota Udaipur has the highest literacy rate and the lowest working population percentage which could probably indicate the impetus required on providing education to people. One could also use a set value for each of these attributes as the ideal value that needs to be achieved under a certain standard of quality of life and compare all the villages to that standard. This again would provide a different set of insights into the major areas of focus and the amount of effort required in achieving the set targets.

Another way of looking at a variable set of attributes and villages is to visualize a straight line from which perpendicular lines emerge as axes for various attributes as shown in figure

5. Hence the initial single straight line can be seen as a line of points that can act as origins for the lines that are perpendicular to them. The villages are again represented as circles with their radius varying proportionately as the population of the area being represented increases. The state of a village in terms of attributes being used is represented as a dashed line that connects the position of the village at every axis. Thus, the dashed line in effect again acts a reference model for comparison between various villages in terms of the attributes being compared.

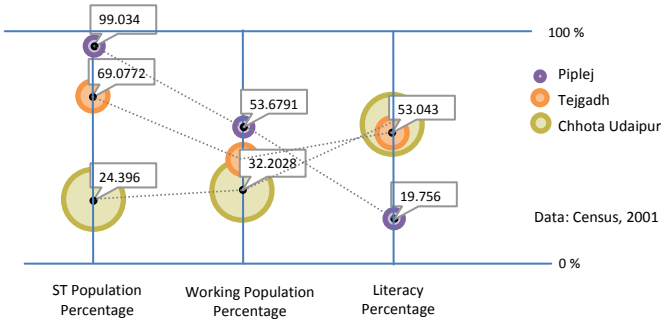

FIGURE 5
VISUAL REPRESENTATION OF COMPARISON BETWEEN THREE VILLAGES IN TERMS OF THE ATTRIBUTES RELATING TO LITERACY, WORKING POPULATION AND ST POPULATION

The attributes being looked at are a product of CENSUS 2001 information. With the amount of data coming into the system one can look at a variety of other attributes for visualization. These attributes would involve *percentage of people suffering from various diseases out of the total population of an area*, percentage of people migrating for work, *percentage of people having access to affordable medical care within an area* and various other factors that determine the quality of life of people in general. These diagrams would also enable cross-comparison and assist in analyzing the actual state of affairs of villages with respect to various attributes and their respective values in terms of percentages.

III. Geographical Information System

Primarily the GIS will be used to identify areas of interest with respect to the scales defined for problems. Given that one can add support for spatial data and perform operations such as area-of-interest queries to a relational database such as *Oracle* or *MySQL*, implementing a GIS would require GIS data in addition to the data being collected by the Field Workers. Presently the system incorporates only the latitudes and longitudes of places and it enables one to effectively pin point villages of *Virtual Earth*.

The following three methodologies can be used to look at information in terms of geographical maps:

- **Latitude and Longitudinal Representation:** Given that one knows precise location of villages in and around Tejgadh in terms of latitude and longitude creating, alerts with respect to certain problems would provide an effective way of looking at data being collected. All that one requires is a scale of evaluating whether a problem requires an alert in a particular area or not. With the semantics of a problem figured out, representation is merely a function of creating certain recognizable shapes with distinct colors to define the area and extent of the problem. Presently the representation of villages in terms of pushpins has been accomplished. The attributes for data representation are chosen and then, represented on the precise location of the area of a map using a pushpin as shown in figure 6.

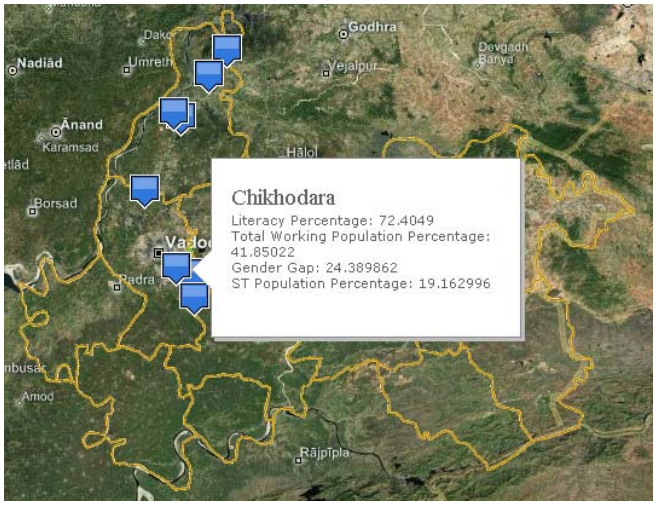

FIGURE 6
PUSHPINS REPRESENTING INFORMATION ON VILLAGES IN TERMS OF ATTRIBUTES THAT ARE SELECTED WITH A MASHUP OF VADODARA DISTRICT IN THE BACKGROUND

- **Overlays:** Regions on *Virtual Earth*[2] could be overlaid with layers of geological maps, rainfall maps, and any other map representing vital information about a region. One may also create an additional layer of marking boundaries of each and every village on the maps in terms of latitudes and longitudes.

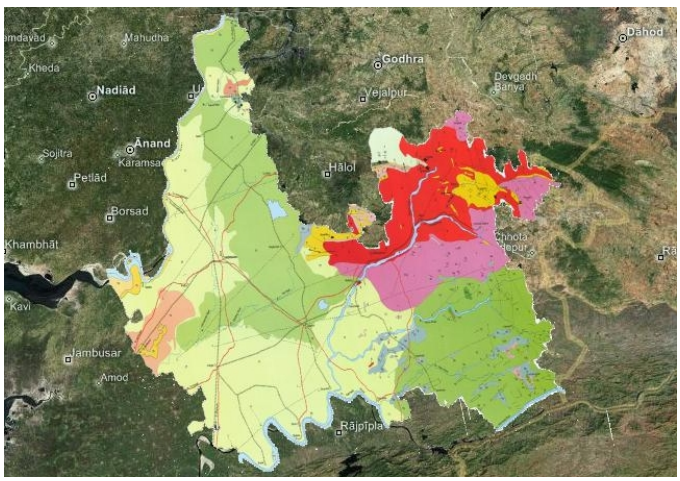

FIGURE 7
OVERLAY OF DISTRICT RESOURCE MAP OF VADODARA ON VIRTUAL EARTH

Once the boundaries are marked, one can color code them in terms of the problems that a particular village primarily faces. The intersection of these overlays would create secondary opportunities of interpreting what problems together add up to affect the quality of life of Tribal people in general. For example, an intersection of areas having people with low income and areas having majority of illiterate people should give a fair estimate of the impact of education in the area and would also suggest places that require establishment of schools providing subsidized education. The overlays of maps on *Virtual Earth* has been accomplished using *MSR Map Cruncher*[3] and a shape file for marking boundaries of each and every village has also been obtained. Currently, attempts are being made to map the shape file on *Virtual Earth* and then mark boundaries using pushpins.

- **Movement Patterns:** Migration patterns of people could be evaluated by connecting information regarding places people generally move to find employment opportunities to the original native villages. This information could be used to identify potential causes of a few diseases such as silicosis. If people migrate to work in limestone quarries then there is a high probability that they might get infected by silicosis. Thus the connection between areas is a reflection upon the health and the occupational patterns of people. Presently the work on this module is under progress and research is going on as to how to create representational polygons for connecting places on a map.

The GIS has been implemented on *Virtual Earth* as it provides the flexibility of creating mashups using static images of maps. The tool that has been used for creating such mashups is *MSR Map Cruncher*. There is ample documentation on using *Virtual Earth SDK*[4] for various representational schemes and it does not require any key or registration for usage.

6. Expected Impact

The solution hopes to provide essential control points or focus areas that need to be looked at while making attempts at developing an area. It provides an estimation of the problems a particular area is facing in terms of the effects that can be observed in representations. The initial usage of the application is primarily to identify regions that require fire-fighting against a problem. With gradual increase of data, one may achieve a holistic approach to policy making by having accurate estimations of the problems that the Tribal Belt faces. Ultimately it will help in evaluating the impact of the policies that have been implemented in the area and evolve as the problems faced by regions change and also altering the way data is managed within the organization.

The solution is expected to be a web-service that implements a GIS plus an engine to evaluate data in other representational formats. Thus it represents information in ways that are more diagrammatic than statistical, which are easier to understand and collate. Ultimately it should serve as a tool that assists in development work with more insights into the problem at hand.

7. Conclusion

The present status of the project involves work that has been done on:

- Collection of data and survey formats from the Academy of Tribal Studies, Tejgadh
- Collection of Maps from BISAG
- Development of Concept Maps and Database Design.
- Creation of dynamic data entry mechanism for surveys
- Creation of two visualization techniques with prototypes in *Flash* and implementation in *JSP*,
- Identification of a GIS platform in Virtual Earth
- Creation of mashups of the district of Vadodara using the maps supplied by BISAG.
- Entry of latitude and longitude data using the maps provided by BISAG into the database.
- Creating pushpins for representing various villages on *Virtual Earth* and representing values based on the selected attributes.

The database schema has been finalized with a proper interaction with the members of the NGO. The prototype for data entry mechanism has been created. Similarly prototypes for representational visualizations and geographic information systems have been created. All these prototypes are stand alone applications that need to integrated into a website that provides access to all this after identifying the user level.

Once the website is in place the representational methodologies could be refined using data digitized for ten sample villages. The future work in the project also involves data collection and training programs for the Field Workers to initiate the data collection methodology. Representations would be re-evaluated and reviewed based on the feedback of the users specified in Section 3. Finally, the project would leave scope for additional representations that can be created later as a different project altogether.

Acknowledgement

I would like to express my sincere gratitude to Prof. Binita Desai for her constant guidance and support throughout the span of the project. I would like to thank Navin Goel for his help during the implementation of the project.

References

[1] Oracle Spatial, Locator and Location-Based Services accessed at *http://www.oracle.com/technology/products/spatial/index.html*

[2] Virtual Earth : Home accessed at *http://www.microsoft.com/virtualearth/default.mspx*

[3] MSR Map Cruncher accessed at *http://research.microsoft.com/mapcruncher/*

[4] Virtual Earth Interactive SDK accessed at *http://dev.live.com/virtualearth/sdk/*