GEORGIA DOT RESEARCH PROJECT 16-23
FINAL REPORT

HR DATA TOOL (HRDT): A MODULAR SYSTEM FOR SUPPORTING GDOT HUMAN RESOURCE PLANNING AND DECISION MAKING

OFFICE OF PERFORMANCE-BASED MANAGEMENT
AND RESEARCH
15 KENNEDY DRIVE
FOREST PARK, GA 30297-2534

TECHNICAL REPORT STANDARD TITLE PAGE

1.Report No.: FHWA-GA-19-1623	2. Government Accession No.:	3. Recipient's Catalog No.:
4. Title and Subtitle: HR Data Tool (HRDT): A Modular System for Supporting GDOT Human Resource Planning and Decision Making		5. Report Date: March 2019 6. Performing Organization Code:
7. Author(s): Dr. J. David Frost, Ph.D., P.E.		8. Performing Organ. Report No.:
9. Performing Organization Name and Address: Georgia Institute of Technology School of Civil & Environmental Engineering 790 Atlantic Drive Atlanta, GA 30318		10. Work Unit No.: 11. Contract or Grant No.: 0015163
12. Sponsoring Agency Name and Address: Georgia Department of Transportation Office of Performance-Based Management and Research 15 Kennedy Drive Forest Park, GA 30297-2534		13. Type of Report and Period Covered: Final; July 2016-March 2019 14. Sponsoring Agency Code:
15. Supplementary Notes: Prepared in cooperation with the U.S. Department of Transportation, Federal Highway Administration		

16. Abstract:
This report presents the concepts and framework behind Human Resources Data Tool (HRDT), a Java-based software package developed specifically for the Georgia Department of Transportation (GDOT) to aid in analysis of Human Resources (HR) data and for use in planning and decision making. This project is built upon the previous JOB SEEKER (Job Shadowing for Employee Engagement through Knowledge and Experience Retention) project (GDOT Research Project No. RP13-12) by applying data mining and data analysis techniques such as network analysis, spatial analysis, and pivot charts and tables, in order to gain a deeper understanding and insight into the existing GDOT HR data.

The software includes seven data modules: (i) Job Shadowing, (ii) Workforce Planning, (iii) Succession Planning, (iv) Cross Training, (v) Training and Development, (vi) Network Analysis, and (vii) Decision Making. These data modules can allow GDOT HR personnel to assess workforce and training needs not only at an individual level, but they can also be used for planning and strategic decision-making at the organizational level.

17. Key Words: Human resources management, data analysis, planning, decision making		18. Distribution Statement:	
19. Security Classification (of this report): Unclassified	20. Security Classification (of this page): Unclassified	21. Number of Pages: 107	22. Price:

Form DOT 1700.7 (8-69)

GDOT Research Project No. RP16-23

Final Report

HR Data Tool (HRDT): A Modular System for Supporting GDOT Human Resource Planning and Decision Making

By
J. David Frost, Ph.D., P.E.
Principal Investigator

Georgia Institute of Technology

Contract with

Georgia Department of Transportation
In Cooperation with
U.S. Department of Transportation
Federal Highway Administration

March 2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

A project titled "JOB SEEKER" (Job Shadowing for Employee Engagement through Knowledge and Experience Retention) was previously performed by the Georgia Institute of Technology for the Georgia Department of Transportation (GDOT Research Project No. RP13-12), and the report was submitted in May 2016 (Report No. FHWA-GA-16-1312). The purpose of that project was to explore how to optimally use the knowledge retention/transfer technique of "job shadowing" as an informal method for knowledge capture and transfer as well as increasing communication and employee engagement, particularly for capturing and disbursing knowledge between the "near-retirement" generation and the "new generation" of workers.

During the course of the JOB SEEKER project, attrition data were provided by GDOT Human Resources (HR) group in the form of a spreadsheet containing the location, working titles (i.e., positions), and years of service of almost 4,000 GDOT employees, as well other attributes of the agency workforce. At that time, it became apparent that GDOT did not have the capability to rigorously analyze the employee data, and that analysis of the existing data could provide valuable insights not only with respect to job shadowing, but also for institutional planning and strategic decision making.

This project builds upon the previous JOB SEEKER project by applying data mining and analysis techniques such as network analysis, spatial analysis, and pivot charts and tables in order to gain a deeper understanding and insight into the existing GDOT employee database. The project has resulted in the creation of an integrated software platform called "HR Data Tool" (HRDT). HRDT is a modular system, coded using Java

programming language that can explore and leverage existing GDOT HR data to assist in organization-wide training, planning and decision-making.

HRDT includes a modular and consistent framework for evaluating various parameters which are important to planning and decision making. The software includes seven data modules:

- Job Shadowing
- Workforce Planning
- Succession Planning
- Cross Training
- Training and Development
- Network Analysis
- Decision Making

These data modules can allow GDOT HR personnel to assess workforce and training needs not only at an individual level, but they can also be used for planning and strategic decision-making at the organizational level. It is anticipated that HRDT will help GDOT to advance from being an information-rich organization to a knowledge-rich one, in turn helping GDOT achieve its stated mission of delivering "a transportation system focused on innovation, safety, sustainability and mobility" and fulfilling its goal to "recruit, train and retain a quality workforce".

ACKNOWLEDGMENTS

A number of individuals provided valuable assistance during the course of this study. They included individuals from various GDOT units who provided critical insights into current practices and experiences, and participated in valuable project-related discussions. In particular, the contributions of Monica Ivey, Brian Robinson, and Supriya Kamatkar throughout the study period are sincerely appreciated. A number of other GDOT employees from various offices also contributed to the work through participation in informal and formal discussions. Their engaged participation is recognized.

1. INTRODUCTION

A project titled "JOB SEEKER" (Job Shadowing for Employee Engagement through Knowledge and Experience Retention) was previously performed by the Georgia Institute of Technology (GDOT Research Project No. RP13-12), and the report was submitted in May 2016 (Report No. FHWA-GA-16-1312). The purpose of that project was to explore how to optimally use the knowledge retention/transfer technique of "job shadowing" as an informal method for knowledge capture and transfer as well as increasing communication and employee engagement, particularly for capturing and disbursing knowledge between the "near-retirement" generation and the "new generation" of workers.

During the course of the JOB SEEKER project, attrition data were provided by GDOT Human Resources (HR) group in the form of a spreadsheet containing the location, working titles (i.e., positions), and years of service of almost 4,000 GDOT employees, as well as other attributes of the agency workforce. At that time, it became apparent that GDOT did not have the capability to rigorously analyze the employee data, and that analysis of the existing data could provide valuable insights not only with respect to job shadowing, but also for institutional planning and strategic decision making.

This project builds on the previous JOB SEEKER project by applying data mining and data analysis techniques such as network analysis, spatial analysis, and pivot charts and tables, in order to gain a deeper understanding and insight into the existing GDOT employee database. The project has resulted in the creation of an integrated software platform called "HR Data Tool" (HRDT). HRDT is a modular system, coded using Java

programming language that can explore and leverage existing GDOT HR data to assist in organization-wide training, planning and decision-making. Figure 1-1 shows the home (landing) page of the HRDT software.

Figure 1-1 – HR Data Tool home page

1.1 PURPOSE

Knowledge and experience within GDOT are both a major investment and a valuable resource. Equally well, these are some of the most vulnerable assets that can be easily impacted or lost. As such, developing strategies to identify potential opportunities for knowledge and talent management, as well as for institutional planning and strategic decision making, are crucial to the success of GDOT operations.

The HRDT project aims to address these issues by utilizing data mining, data analysis, and network analysis techniques to address knowledge loss and operational challenges before they become problematic for GDOT. It also serves as a planning tool for HR personnel to advise GDOT leadership with respect to critical personnel planning and training needs across the organization. In this regard, HRDT can be seen as an integral component of GDOT's mission statement of delivering "a transportation system focused on innovation, safety, sustainability and mobility" and the associated stated goal to "recruit, train and retain a quality workforce".

1.2 METHODOLOGY

After project initiation, several meetings were held with GDOT HR personnel to discuss the proposed approach and obtain feedback with regard to the research and program specifics. A framework was then developed for HRDT to satisfy the project goals.

The framework includes tools for identification of relevant employee factors from the GDOT provided employee database, including systematic identification of vacancy risk (i.e., how close an employee is to retirement, based on his/her age and/or tenure) and other factors, criticality and uniqueness of each position within the organization, and connectivity of the individuals within the organization. The framework was programmed in Java programming language, and a "Beta" version of the software was delivered to GDOT HR personnel for testing and evaluation. Suggestions made by GDOT personnel were then incorporated to create the final version of HRDT. In addition, a User's Manual has been developed for future reference (see Appendix A) and a Training Module to

assist with introducing the software to future cohorts of GDOT personnel (see Appendix B).

1.3 Report Organization

The project report is organized as follows:

- Section 1 provides a brief introduction, as well as purpose and methodology for the research study.
- Section 2 contains relevant background and overview on data mining, pivot table/chart and spatial analysis, multi-variate analysis, and network analysis, particularly as they relate to HR data.
- Section 3 provides a description of the methodology and theory associated with the development of the HR Data Tool (HRDT) software.
- Section 4 contains project conclusions.
- Section 5 contains report references.

Appendix A contains the User's Manual for HRDT. Appendix B contains the Training Module presentation slide deck.

2. BACKGROUND & OVERVIEW

2.1 DATA MINING

Data mining can be described as the process of discovering patterns and extracting knowledge from large amounts of data. Organizations are often data rich but information poor, because not enough is done to extract knowledge from existing databases and other sources of data (Han et al., 2012). With regard to human resources management, data mining is an integral part of knowledge management (Silwattananusarn and Tuamsuk, 2012). Data mining can be used for talent management, scenario forecasting, human talent predictions, and strategic decision making (Ranjan et al., 2008; Jantan et al., 2010; Sadath, 2013).

The data mining process is an iterative sequence that involves the following (Han et al. 2012); the first four steps involve pre-processing of the data to prepare it for mining:

- Data cleaning (to remove noise and inconsistent data)
- Data integration (combining multiple databases if applicable)
- Data selection (retrieval of relevant data from the database)
- Data transformation (transformation and consolidation of data into forms appropriate for mining)
- Data mining (application of intelligent methods to extract data patterns)
- Pattern evaluation (to identify patterns representing knowledge)
- Knowledge presentation

Data mining adopts techniques from many fields (see Figure 2-1). Several of the techniques shown in this figure were utilized in this project, including statistics, information retrieval, algorithms and visualization.

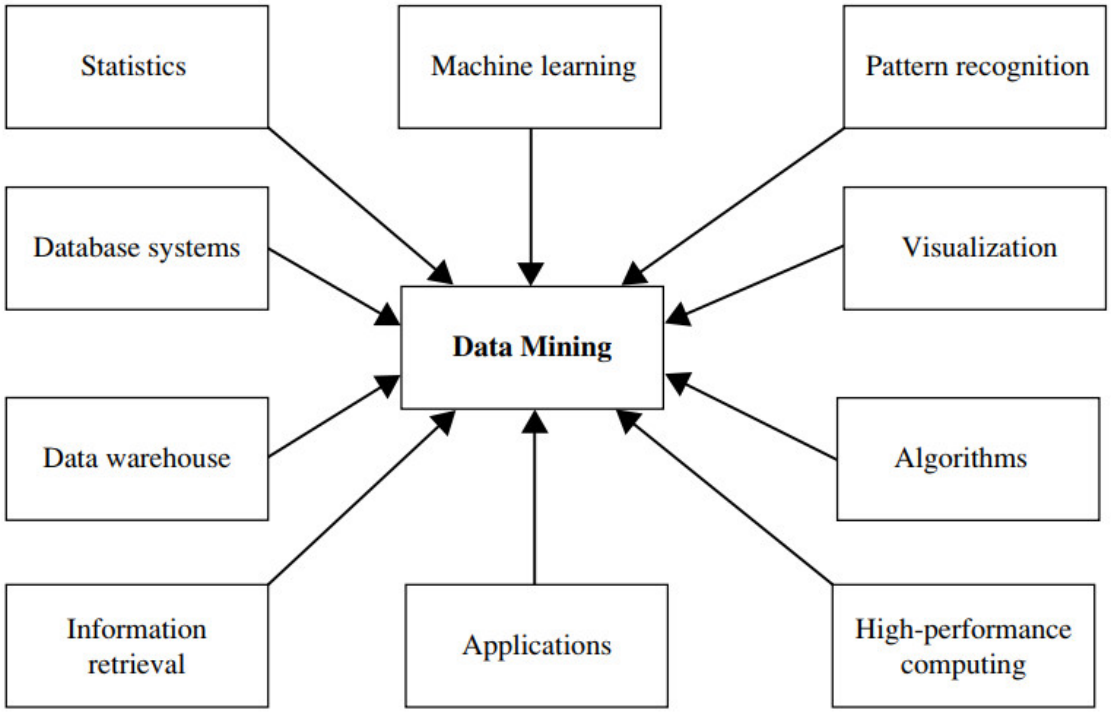


Figure 2-1 – Domains relevant to data mining (from Han et al. 2012)

2.2 DATA ANALYSIS

2.2.1 PIVOT TABLE/CHART & SPATIAL ANALYSIS

A pivot table is a dynamic data summarization and analysis tool which can allow sorting, counting, averaging, and performing many other mathematical and statistical operations on data extracted from a database. Similarly, a pivot chart is a data analysis tool that enables the visualization of the results of a pivot table. The combination of pivot

tables and charts allow for dynamic analysis and visualization of the data, once relevant data have been extracted from a database. In the context of human resource management, the application of pivot table and chart techniques can provide a powerful tool that allows the identification of critical positions within the organization with respect to attrition risk. This in turn, can be used for succession planning purposes and/or for facilitating knowledge capture/transfer.

As an example, Figure 2-2 shows the results of pivot table and chart analysis looking at the District Engineer position of an organization. It can be seen about half of the district engineers are expected to retire in the next 2 years. Information such as this can allow the decision makers at the organization to ensure that appropriate measures are taken to minimize disruptions related to potential attrition and assure continuity.

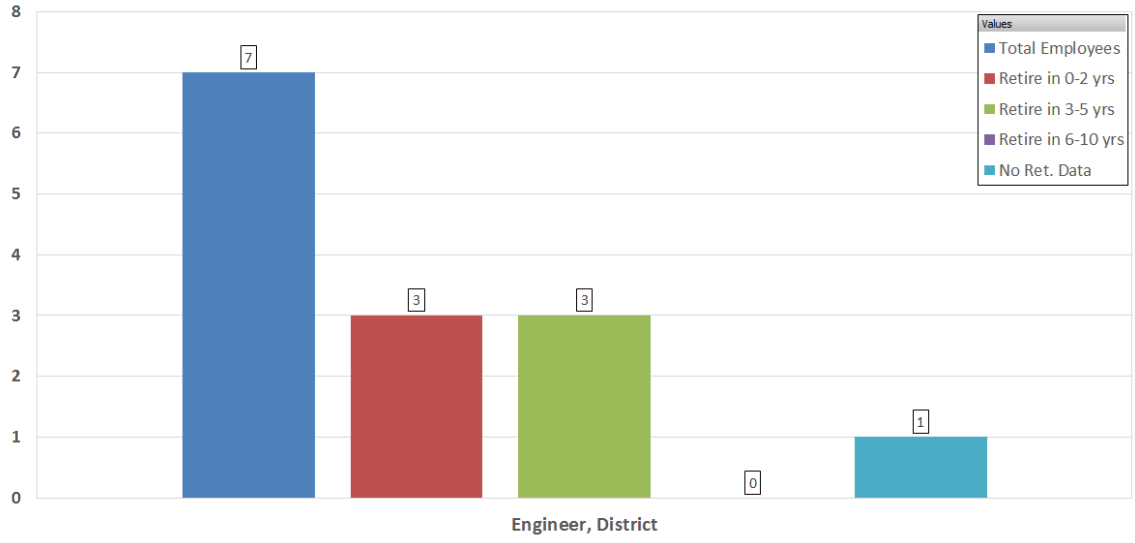


Figure 2-2 – An example of pivot table/chart analysis

Additionally, in the case of organizations with many offices across a state, country or even the world, it is desirable to have an understanding of the spatial distribution of employee data across all of the offices. In this regard, the use of spatial analysis techniques can allow for management, analysis, and visualization of employee data spatially and dynamically to understand patterns, trends, and relationships. This can be accomplished by integrating the results from data analysis with spatial analysis tools such as ArcGIS, as will be discussed later in this report.

2.2.2 MULTI-VARIATE ANALYSIS

Generally speaking, multi-variate analysis refers to statistical techniques considering two or more variables. In the context of this report, multi-variate analysis was used to calculate scores for Knowledge Loss Risk (which is a function of four different factors), Mentor Evaluation, and Protégé Evaluation, as will be discussed in more detail later in this report.

2.2.3 NETWORK ANALYSIS

Network analysis is a multi-disciplinary field which seeks to predict and examine the interaction and effects of objects (nodes) inside a network which are connected to each other through a predefined relationship among the nodes. Network analysis is a subset of graph analysis that involves examining graphs as a representation of symmetric and asymmetric relations (directed and undirected graphs) between objects.

The concept of network analysis is employed in many fields, including physical and social sciences. A lot of valuable information could be inferred from the relationship between humans and communication within the structure of a network, such as finding the most influential person in a network and how fast information can diffuse throughout the whole network. In this regard, the application of such network analysis techniques to an organization like GDOT can allow for the identification of critical knowledge and connectivity between individuals within the organization, which in turn can have important implications for organization-wide planning and strategic decision making .

In the context of an organization, a network can simply be defined as an interconnected group of people or things (such as computers, operations, etc.). Information can move around organizations through hard networks (infrastructure-dependent) as well as soft networks (informal and typically based on social interaction). In organizations consisting of multiple sub-units, professional networks and the relationships between the network members can be especially important in determining the flow and sharing of data and knowledge through the network (Hansen, 1999; Reagans & McEvily, 2003; Hansen, Mors, & Lovas, 2005).

Relative importance, or centrality, of the members also has important implications for professional networks. One common technique to evaluate the centrality of a member in a network is referred to as "betweenness" (Freeman, 1978/1979; Butts, 2008). Betweenness centrality is an indicator of a member's centrality in a network, with high betweenness individuals acting as "bridges" between different groups that may otherwise be loosely connected (see Figure 2-3). These individuals tend to have a large influence on the

sharing and transfer of knowledge through a network, assuming that the transfer takes place along the shortest paths associated with a given network member (Barthelemy, 2004; Butts, 2008). An example application of betweenness centrality is a network representation of a classical organizational chart. By representing the organizational chart as a network, it is possible not only to see who reports to who, but also to visualize the connectivity of the individuals within the network.

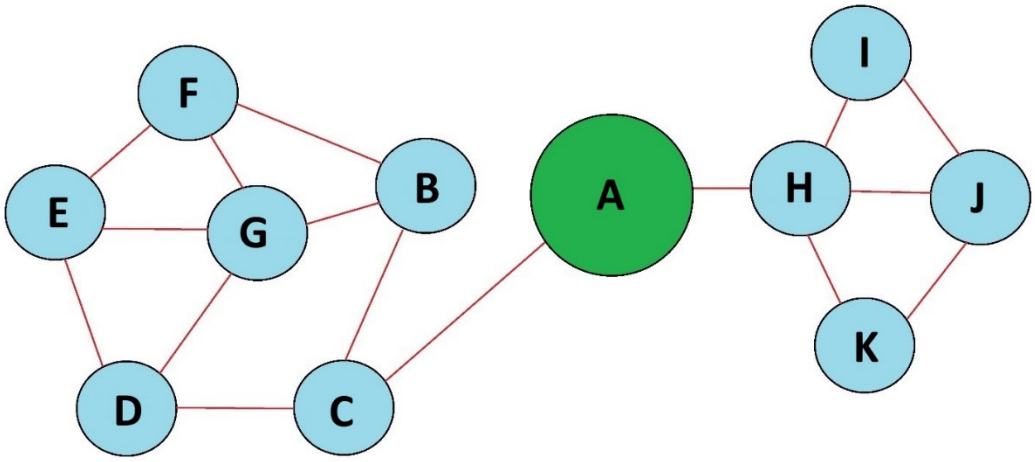


Figure 2-3 – An example of betweenness centrality; node A has the highest centrality value as it acts as a bridge which connects all the cluster of left nodes to the cluster of right nodes.

Betweenness centrality only takes into account the immediate ties that a member has in a network. There might be situations where one member "might be tied to a large number of others, but those others might be rather disconnected from the network as a whole. In a case like this, the node (member or employee) could have high centrality, but only in a local neighborhood" (Hanneman and Riddle, 2005). Closeness centrality is a technique to overcome this limitation. In particular, PageRank centrality is a form of

closeness centrality, which considers three distinct factors that determine the centrality of a node: (i) the number of links it receives, (ii) the link propensity of the linkers, and (iii) the centrality of the linkers. PageRank centrality was developed by Google founders (Brin and Page, 1998) to more efficiently rank web search engine results. The algoritm considers not only the number of the links in a network, but also the quality of the links.

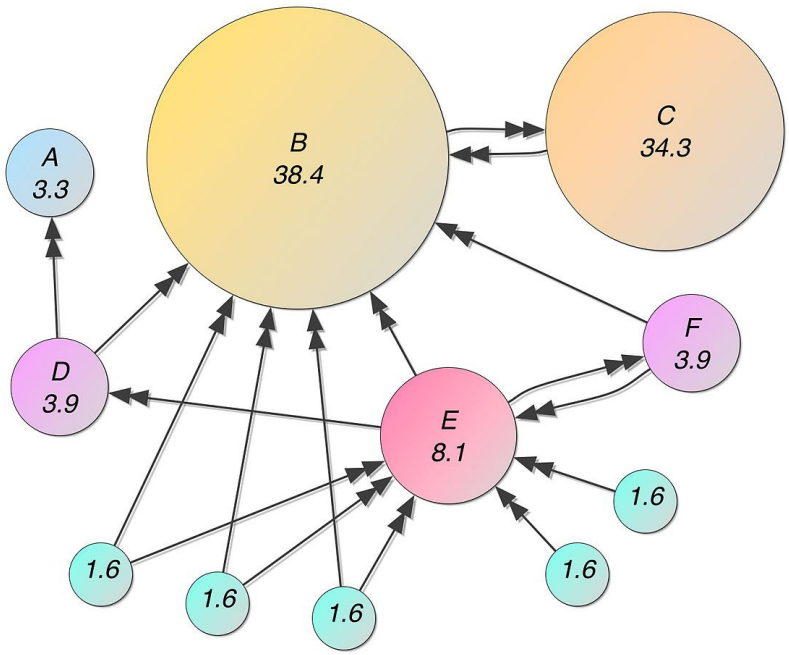


Figure 2-4 – An example of PageRank centrality; node C has a higher rank (values are percentages) than node E because it is linked to node B, even though there are fewer links to C (from Wikimedia Commons, 2019).

Another important concept in network analysis is called cosine similarity. Cosine similarity is used to determine likeness between two vectors. For example, a vector can consist of certain employee attributes, and cosine similarity can be used to calculate the degree of similarity between two employees in a network. A value of 0 indicates that there are no similaririties between vectors (i.e., the angle between the two vectors is 90 degrees), while a value of 1 indicates the vectors are identical (i.e., the angle between the two vectors is 0 degrees).

2.2.4 SYNOPSIS

In summary, the combination of techniques such as data mining, pivot table/charts and spatial analysis, multi-variate analysis and network analysis can be used to identify critical knowledge holders within the organization, to prioritize amongst potential candidates (for any given position) by considering the position of the node (person) in a network and its connectivity, and to provide GDOT leadership with guidance for organization-wide planning, training, development, and strategic decision making. Network analysis can allow transformation from an individual scale (i.e., one-on-one) to an organizational scale with regard to human resources management.

3. GDOT HR DATA TOOL

HRDT was developed using the existing employee database as the input, and takes into account the characteristics and needs of the GDOT organization. The HRDT software components are shown in Figure 3-1. The details of the system components are discussed further in the subsequent sections.

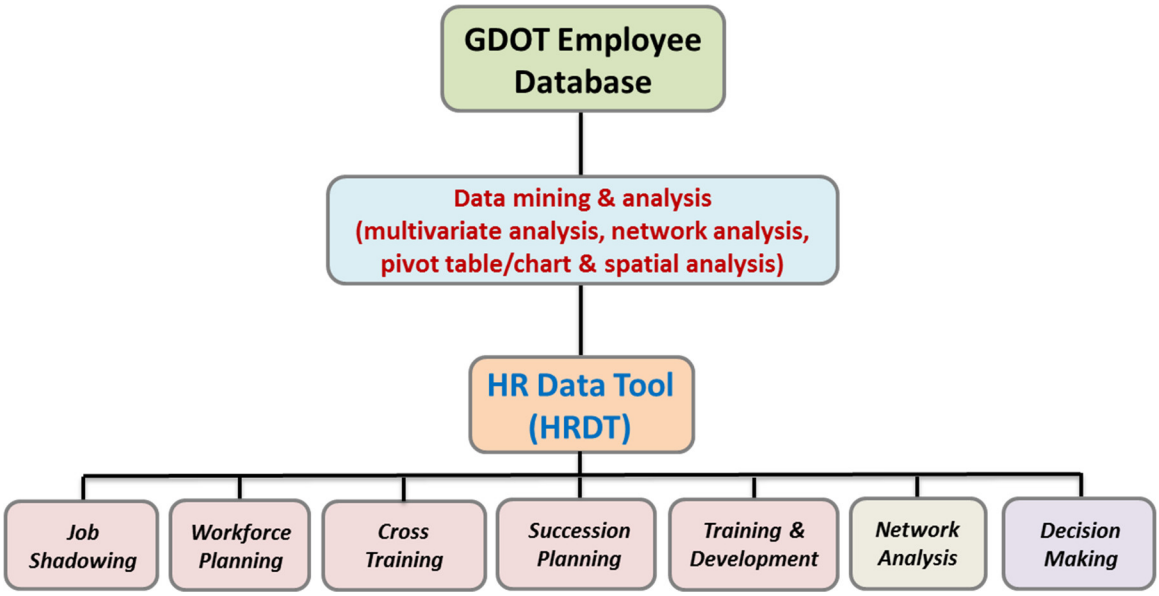


Figure 3-1 – HRDT software components

3.1 Database Structure & Data Mining

The provided employee database consists of 35 columns and approximately 4,000 rows, and includes information such as employee's name, job title, department, age, overall start date, start date at current position, supervisor name, among other attributes.

The physical location (district and area) of the employees had to be extracted from the data as it was not explicitly given, by using a text string search algorithm to identify the district and area to which the employee is assigned from his/her department. The employees were then assigned to the physical address associated with each district and area office. It should be noted that the GDOT Atlanta headquarters is denoted as District 0-Area 0 (see Table 3-1).

Table 3-1 – GDOT District and Area locations

District-Area	Name	Address
0-0	GDOT Headquarters	600 West Peachtree NW, Atlanta, GA 30308
1-0	Gainesville (Main)	2505 Athens Hwy SE, Gainesville, GA 30507
1-1	Gainesville (Area 1)	2594 Gillsville Hwy, Gainesville, GA 30507
1-2	Athens	450 Old Hull Rd, Athens, GA 30601
1-3	Carnesville	301 Conger Rd, Carnesville, GA 30521
1-4	Cleveland	942 Albert Reid Rd, Cleveland, GA 30528
2-0	Tenille (Main)	643 Hwy 15 S, Tennille, GA 31089
2-1	Milledgeville (Area 1)	161 Blandy Rd, Milledgeville, GA 31061
2-2	Dublin	2003 US Hwy 441 S, Dublin, GA 31021
2-3	Louisville	2791 US Hwy 1 N, Louisville, GA 30434
2-4	Augusta	4260 Frontage Rd, Augusta, GA 30909
2-5	Madison	1570 Bethany Rd, Madison, GA 30650
3-0	Thomaston (Main)	115 Transportation Blvd, Thomaston, GA 30286
3-1	Thomaston (Area 1)	101 Transportation Blvd, Thomaston, GA 30286
3-2	Columbus	3600 Schatulga Rd, Columbus, GA 31907
3-3	Perry	200 Julianne St, Perry, GA 31069

Table 3-1 (cont.) – GDOT District and Area locations

District-Area	Name	Address
3-4	Macon	4499 Riverside Dr, Macon, GA 31210
3-5	LaGrange	1107 Hogansville Rd, LaGrange, GA 30241
4-0	Tifton (Main)	710 West 2nd St, Tifton, GA 31794
4-1	Valdosta (Area 1)	1411 Madison Hwy, Valdosta, GA 31601
4-2	Douglas	1835 S Peterson Ave, Douglas, GA 31535
4-3	Donalsonville	734 W Crawford St, Donalsonville, GA 39845
4-4	Moultrie	120 Veterans Pkwy N, Moultrie, GA 31788
4-5	Albany	2060 Newton Rd, Albany, GA 31701
5-0	Jesup (Main)	204 North Highway 301, Jesup, GA 31546
5-1	Baxley (Area 1)	740 Oakdale Cir, Baxley, GA 31513
5-2	Waycross	104 N Nichols St, Waycross, GA 31502
5-3	Brunswick	128 Public Safety Blvd, Brunswick, GA 31525
5-4	Statesboro	17213 US Hwy 301 N, Statesboro, GA 30458
5-5	Savannah	630 West Boundary St, Savannah, GA 31401
6-0	Cartersville (Main)	500 Joe Frank Harris Pkwy, Cartersville, GA 30120
6-1	Cartersville (Area 1)	874 Peeples Valley Rd NW, Cartersville, GA 30120
6-2	Dalton	1313 North Tibbs Rd, Dalton, GA 30720
6-3	Buchanan	4323 US Hwy 27, Buchanan, GA 30113
6-4	Rome	533 East 20th St, Rome, GA 30161
7-0	Chamblee (Main)	5025 New Peachtree Rd, Chamblee, GA 30341
7-1	Chamblee (Area 1)	5025 New Peachtree Rd, Chamblee, GA 30341
7-2	Marietta	1296 Kennestone Cir, Marietta, GA 30066
7-3	College Park	4125 Roosevelt Hwy, College Park, GA 30349

Additional knowledge extracted from the database includes the projected retirement of each employee (using a formula provided by GDOT which considers the age and the tenure of the employee), the tenure at current position, as well as the managerial reporting structure within the organization (i.e., who reports to who). In this regard, one of the most powerful aspects of HRDT is that it can automatically create an organizational chart for 56 pre-defined departmental units (as obtained from the Master Organizational Chart), going as high as the level of the commissioner. These departmental units are summarized in Table 3-2:

Table 3-2 – GDOT departmental units used in HRDT

Name	Name
Office of Planning	Office of Utilities
Office of Human Resources	Office of Traffic Operations
Office of Legal Services	Office of Maintenance
Office of Equal Employment Opportunity (EEO)	Office of Innovative Delivery
Office of Strategic Communications	Office of Program Delivery
Office of Procurement	Office of Program Control
Information Technology	Office of Engineering Services
Office of Application Support	Office of Transportation Investment Act (TIA)
Office of Infrastructure	Office of Performance-Based Management and Research
Office of Local Grants	Office of Budget Services
District 1	Office of Financial Management
District 2	Office of General Accounting
District 3	General Counsel Division of Administration
District 4	Division of Local Grants
District 5	Division of Engineering

Table 3-2 (cont.) – GDOT departmental units used in HRDT

Name	Name
District 6	Division of Intermodal
District 7	Division of Construction
Office of Equipment and Facilities Management	Division of Permits and Operations
Office of Environmental Services	Division of Public-Private Partnerships (P3)
Office of Roadway Design	Program Delivery
Office of Bridge Design	Division of Finance
Office of Right of Way	Deputy Commissioner
Office of Design Policy and Support	Chief Engineer
Office of Intermodal	Treasurer
Office of Materials and Testing	Division of Planning
Office of Construction	Office of Audits
Office of Bidding Administration	Government and Legislative Relations
Office of Transportation Data	Commissioner

It is very important to note that HRDT relies on a database generated by GDOT Information Technology (IT) personnel as input. The latest database provided was an .XLS (Microsoft Excel 97-2003) worksheet file created using a Structured Query Language (SQL) inquiry with the statement as given in Table 3-3. In addition, two new columns were added so that the "PE" (PE = Professional Engineer) and "EIT" (EIT = Engineer-in-Training) (Columns AL and AM, respectively) information can be added manually. **It is crucial to update the database such that it follows strictly the same column format as the current database. Any deviations from the current format will result in HRDT not functioning as intended.**

Table 3-3 – SQL statement used by GDOT to generate the HRDT input database

```
    select h.DEPTID, h.DEPTNAME , dc.CHARGE_DEPTID,
dc.CHARGE_DEPTID_DESCR, dc.FUND_SRC, dc.FUND_SRC_DESCR,
dc.PROGRAM, dc.PROGRAM_DESCR,

        h.emplid, h.LAST_NAME, h.FIRST_NAME , h.SEX as GENDER, ed.AGE,
ed.ETHNIC_GROUP_CD, ed.ETHNIC_GROUP,   ed.YEARS_OF_SERVICE as
gdot_years_of_service ,   h.JOBCODE, h.JOBTITLE, h.JOB_ENTRY_DT,

        h.POSITION_ENTRY_DT, h.POSITION_NBR, h.FLSA_STATUS,
h.GRADE , h.ANNUAL_RT, h.PAYGROUP, h.ORIG_HIRE_DT,  h.REHIRE_DT,
h.REG_TEMP, h.DX_DRUG_TEST, h.COUNTY, h.EEO_CLASS,

        h.REPORTS_TO as SUPERVISOR_POSITION_NBR,

        hm.EMPLID as Supervisor_EmplID, hm.LAST_NAME as
SUPERVISOR_LAST_NAME, hm.FIRST_NAME as
SUPERVISOR_FIRST_NAME,  hm.JOBCODE  as SUPERVISOR_JOBCODE,
hm.JOBTITLE as SUPERVISOR_JOBTITLE

      from SAO.VW_HCM_DOT_EMPLOYEE_MV h,

        SAO.VW_HCM_DOT_EMPLOYEE_MV hm,

        gdot_hr.employee_datamart  ed,

        GDOT_BUDGET.VW_DEPARTMENT_CHARGES_CURRENT dc

     where h.REPORTS_TO = hm.POSITION_NBR (+)

      and h.DEPTID = dc.DEPTID (+)

      and h.emplid = ed.gdot_employee_id (+)

      order by h.DEPTID, h.emplid
```

3.2 NETWORK ANALYSIS

Previously in the JOB SEEKER project, four factors were defined to perform Knowledge Loss Risk (KLR) assessment, as well as Mentor and Protégé evaluation. These four factors are summarized in Table 3-4 below:

Table 3-4 – Summary of factors used in JOB SEEKER

Factor	KLR Assessment	Mentor Evaluation	Protégé Evaluation
1	Vacancy Risk	Willingness	Past Performance
2	Uniqueness	KLR	Willingness
3	Criticality	Time Period	Time Availability
4	Resource Availability	Time Availability	Location

These dimensions were then assigned a score of either 1, 2 or 3 (with 1 being lowest and 3 being highest), the scores were weighted using the Rank Order Centroid method, and an overall score was then calculated for each factor.

In HRDT, the same general framework has been adopted for consistency, and "Vacancy Risk" (based on the employee's anticipated years to retirement) and "Resource Availability" are calculated the same as in JOB SEEKER. However, while uniqueness and criticality were previously user-assigned values, HRDT performs an assessment based on the data extracted from the database and using network analysis techniques to

provide a first-order estimate of both uniqueness and criticality of each employee. Further, an additional factor termed the "Absence Impact" has been defined for use in a new metric called the "Position Evaluation". Absence Impact can be defined as how important a particular employee and their position is, based on their centrality in the organization. Lastly, additional network variables have been defined for use in candidate searching (for the Workforce Planning and Succession Planning modules) and for trainer and trainee selection (for the Cross Training module). The following sections describe the techniques used to quantify uniqueness, criticality, absence impact, and the additional network variables used in HRDT.

3.2.1 UNIQUENESS ASSESSMENT

Consistent with the ranking scheme adopted in the JOB SEEKER project, there are three possible values for uniqueness of an employee:

1 – Non-unique 2 – Moderately Unique 3 – Unique

Uniqueness can be defined at two different organizational levels: (i) Local, meaning it is assessed at the departmental unit level (for each of the 58 departmental units considered), and (ii) Global, meaning it is assessed at the level of the entire organization. Further, uniqueness can be assessed at two different operational levels: (i) at the basic level, using the working title as a simple indicator, or (ii) at the network level, using the working title as well as the position level of an employee.

In this regard, the four different possible definitions of uniqueness can be given as follows:

- Basic local uniqueness (UQ_{L1}): the inverse of the number of employees with the same working title in the same departmental unit. For example, if there are five employees in a given departmental unit with the title "Engineer", the basic local uniqueness of Engineer in that departmental unit can be calculated as 1/5.
- Network local uniqueness (UQ_{L2}): the inverse of the number of employees with the same working title AND the same position level in the same departmental unit. For example, if there are five employees in a given departmental unit with the title "Engineer", and three of them are at a "Junior" position level, the network local uniqueness of Junior Engineer can be calculated as 1/3.
- Basic global uniqueness (UQ_{G1}): the inverse of the number of employees with the same working title in the entire organization. For example, if there are 10 employees in the entire organization with the title "Accountant", the basic global uniqueness of Accountant in the entire organization can be calculated as 1/10.
- Network global uniqueness (UQ_{G2}): the inverse of the number of employees with the same working title AND the same position level in the entire organization. For example, if there are 10 employees in the entire organization with the title "Accountant", and five of them are at a "Senior" position level, the network global uniqueness of Senior Accountant in the entire organization can be calculated as 1/5.

It was assumed that if uniqueness is less than a value of 0.25 (i.e., greater than 1 in 4), there is enough redundancy for the employee to be considered non-unique and hence assigned a uniqueness value of 1. If uniqueness was greater than or equal to 0.25 but less than 0.5 (i.e., greater than 1 in 2), the employee was assigned a uniqueness value of 2. Lastly, if uniqueness was 0.5 or greater, the employee was assigned a uniqueness value of 3, indicating the employee is unique.

3.2.2 CRITICALITY ASSESSMENT

Consistent with the ranking scheme adopted in the JOB SEEKER project, there are three possible values for criticality of an employee:

1 – Non-Critical 2 – Moderately Critical 3 – Critical

A simple first-order assessment of criticality of an employee can be made based solely on the information contained in his/her job title. The basic assumption is that employees classified as Director, Manager or Supervisor have a higher level of criticality. Employees classified as Foreman and Superintendent are considered critical if the person has been at the job for at least two years. In this regard, the following criteria were used to make a basic criticality assessment (CR-1):

- CR-1 = 3 if "Working Title" in employee database contains any of the following:

  Dir [Director], Mgr [Manager], Spv [Supervisor]
- CR-1 = 3 if "Working Title" contains any of the following:

  Foreman, Superintendent AND tenure at current position ≥ 2 years
- CR-1 = 2 if "Working Title" contains any of the following:

  Foreman, Superintendent
- CR-1 = 1 for all else

A more rigorous assessment of criticality of an employee can be made using the network analysis concept called PageRank centrality. The general equation for the Page Rank scoring can be expressed as follows:

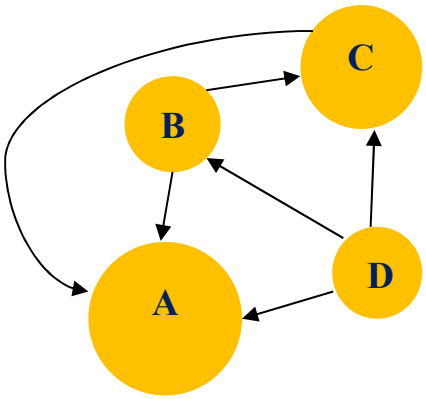
$$PR(\quad) = \sum_{n \in E_m} \frac{PR(n)}{L(n)}$$

This equation shows that Page Rank (PR) for the employee *m* is dependent on PR for any employee *n* in the set E_m, and the number of outgoing edges *L* for employee *n*. This set contains all employees connected to employee *m*. For example, assuming there are 4 employees (A, B, C, and D) in a company, initially all employees will get the same value of PR, which in this example will be 0.25 (considering a probability distribution between 0 and 1). PageRank algorithm is an iterative process. The PR value of an employee will be equally transferred to the neighbor employees upon a new iteration. For example, if all

employees B, C, and D report to employee A (Employees B, C and D are not connected to each other), the PR for Employee A would be:

$$PR(A) = PR(B) + PR(C) + PR(D) = 0.75$$

On the other hand, if Employees A, B, C, and D have the following structure, the PR for employee A upon first iteration would be:



$$PR(A) = \frac{PR(B)}{2} + \frac{PR(C)}{1} + \frac{PR(D)}{3} = 0.458$$

$$PR(B) = \frac{PR(A)}{3} + \frac{PR(D)}{3} = 0.167$$

$$PR(C) = \frac{PR(A)}{3} + \frac{PR(B)}{2} + \frac{PR(D)}{3} = 0.292$$

$$PR(D) = \frac{PR(A)}{3} = .083$$

In the above example, Employee A is considered as a "dangling node" and its effects are equally distributed to other employee nodes in order to have a probability distribution definition over the results of PR values for all the employees. In order to avoid the occurrence of any sinks (dangling nodes), a damping factor is added to the general equation as follows:

$$PR(\quad) = \frac{1-d}{N} \sum_{n \in E_m} \frac{PR(n)}{L(n)}$$

Where *N* is the total number of the nodes and *d* is the damping factor which is usually taken as 0.85.

Another network based method for assessment of criticality of an employee is cosine similarity, which can be generally expressed as follows:

$$Cosine\ Similarity = \cos(\theta) = \frac{A.B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} {A_i}^2}\sqrt{\sum_{i=1}^{n} {B_i}^2}}$$

where, A_i , B_i are attributes of vector *A* (Employee A) and *B* (Employee B). A cosine similarity with zero value represents orthogonality, decorrelation or independency of data meaning that two vectors are not similar to each other. Alternatively, a cosine similarity of 1.0 indicates that two vectors are exactly the same.

To assess criticality, cosine similarity calculations were performed considering both the employee uniqueness (across four dimensions of uniqueness as defined in Section 3.2.1) and Page Rank centrality performed at both the departmental unit (local) level (PR_L) and the organizational (global) level (PR_G). In this approach, six different indicators are considered in order to measure distance similarity among different instances containing more quantitative information. As different features in one instance can have different weights and scales, data over each feature is normalized to avoid any type of bias toward a specific feature in a vector. Lastly, because employees with higher independence relative to their peers can be assumed to have higher criticality, the calculated cosine similarity value was subtracted from unity to calculate "cosine dissimilarity" in order to identify the employee with higher scores, and therefore, higher criticality.

Based on an evaluation of the histogram of the calculated values, it was assumed that employees with a cosine dissimilarity value less than that corresponding to the 70^{th} percentile are considered non-critical, and hence assigned a criticality value of 1. Those employees with cosine dissimilarity values corresponding to between the 70^{th} and 90^{th} percentile are assigned a criticality value of 2. Lastly, if the cosine dissimilarity value falls into the 90^{th} percentile or above, the employee is assigned a criticality value of 3, indicating the employee is critical.

3.2.3 ABSENCE IMPACT ASSESSMENT

As previously discussed, Absence Impact (AI) is a first-order estimate of how important a particular employee and his/her position is, based on his/her centrality in the organization. There are three possible values for the AI of an employee:

1 – Low 2 – Medium 3 – High

The concept of betweenness centrality was used to quantify AI. Betweenness is a centrality measurement based on the shortest path analysis. For every two employees in a network, there is at least one shortest path, where this path is defined as the minimum number of nodes that is needed to be traversed in order to reach from one node to another one, or if there is a weight for each edge, it is defined as the shortest path from one node to another node which leads to the lowest weight.

The betweenness centrality (*BC*) measurement is defined according to the following equation:

$$BC(\ \ ) = \sum_{m \neq n \neq t} \frac{\sigma_{nt}(\ \ )}{\sigma_{nt}}$$

where σ_{nt} is the total number of shortest path from Employee *n* to Employee *t* in the whole graph. $\sigma_{nt}(\ \ )$ is the number of shortest paths passing through employee *m*. Betweenness centrality is calculated at the departmental unit (local) level, because a calculation at the organizational (global) level places too much emphasis on the persons who are located higher in the organization structure from a hierarchical perspective.

Based on an evaluation of the histogram of the calculated values, it was assumed that employees with a betweenness centrality value less than that corresponding to the 70th percentile are considered non-essential, and hence assigned an AI value of 1. Those employees with betweenness centrality values corresponding to between the 70th and 90th percentile are assigned an AI value of 2. Lastly, if the betweenness centrality value falls into the 90th percentile or above, the employee is assigned an AI value of 3, indicating the employee is essential.

3.2.4 ADDITIONAL NETWORK ASSESSMENT

The following additional variables are calculated automatically via network analysis in HRDT:

- Position level: this refers to the level of the employee in the hierarchical sense within the organizational chart. For example, at the departmental unit (local) level, the departmental unit head is considered to be Level 1, those reporting directly to the departmental unit head are considered to be Level 2, and so on. At the organizational (global) level, the commissioner of GDOT is considered to be Level 1, and those reporting directly to the commissioner are considered to be Level 2, and so on. By default, candidates at the same level as the employee are assigned a score of 3, those one level below are assigned a score of 2, and those two levels or greater below are assigned a score of 1.

- Centrality: this refers to the centrality of the employee based on the cosine dissimilarity measure, as discussed in Section 3.2.2.
- Tenure at Current Position: this refers to the number of years the employee has been at his/her current position. This value is extracted from the employee database. By default, employees who have been at their current position for less than one year are assigned a score of 1, at least one but less than two years are assigned a score of 2, and two or more years are assigned a score of 3.

3.3 MULTI-VARIATE ANALYSIS

In HRDT, the user has the option to select a "basic" or a "network" definition of uniqueness and criticality as described in the previous section. Additionally, there are some user-defined variables that must be selected depending on the module being used. These user-defined variables and suggested scoring criteria are summarized below.

3.3.1 USER-DEFINED VARIABLES

The following user-defined variables are required in HRDT:

- Skill Set: these refer to relevant professional skills and certifications; those with a low level of relevant skills should be assigned a score of 1, those with a moderate level of relevant skills should be assigned a score of 2 (this is also the default value), and those with a high level of relevant skills should be assigned a score of 3. For example, a person applying to become a heavy equipment operator would be considered to have a high skill set if he/she has the required certifications to operate various heavy machinery, moderate skill set if he/she is certified to

operate only one or a few specific types of heavy equipment, and a low skill set if he/she does not have the relevant skills and certifications required for the position.

- Willingness / Attitude: this refers to the willingness / attitude of a potential candidate (in Workforce Planning and Succession Planning modules), and also of either a trainer or trainee (in Cross Training Module and Job Shadowing Module). Those with a low level of willingness should be assigned a score of 1, those with a moderate level of willingness should be assigned a score of 2, and those with a high level of willingness should be assigned a score of 3.
- Performance Score: this refers to the past performance of a candidate that can be based on annual reviews (if available for internal candidates), or based on resume (for external candidates). Those with low perceived performance should be assigned a score of 1, those with moderate perceived performance should be assigned a score of 2, and those with high perceived performance should be assigned a score of 3.
- Time Period: this refers to the time period that a potential mentor has available for participation in a job shadowing program, as the success of a job shadowing program is related to the program's duration. A score of 1 should be assigned if there is less than 3 months available, 2 should be assigned if there is 3 to 6 months available, and 3 should be assigned if there is more than 6 months available.
- Time Availability: this is different from the time period factor, and refers to time available as a percentage of total time that a potential mentor has for participation in a job shadowing or cross-training program. For example, a potential mentor who has only 1 or 2 hours per week available may not be as effective as someone

who has 6 to 8 hours per week or more. A score of 1 should be assigned if 4 to 8 hours per week are available, 2 should be assigned if 8 to 16 hours per week are available, and 3 should be assigned if more than 16 hours per week are available.

- Resource Availability: this refers to the fact that there may be budget and/or time constraints within the organization. A score of 1 should be assigned if there is little to no organizational support, 2 should be assigned if there is some organizational support, and 3 should be assigned if there is full organizational support.
- Location: this refers to whether or not a potential mentor and protégé are located in close proximity for a job shadowing program, as the success of a job shadowing program is related to the participants' proximity to each other. A score of 1 should be assigned if the participants are not co-located and there is no Information Technology (IT) infrastructure in place to facilitate communication (e.g., video-conferences, desktop sharing, etc.), 2 should be assigned if the participants are not co-located but there is IT infrastructure, and 3 should be assigned if the participants are co-located.
- Leadership: this refers to a potential candidate's perceived leadership skills for the purposes of Workforce Planning and Succession Planning. Those with low level of perceived leadership should be assigned a score of 1, those with moderate level of perceived leadership should be assigned a score of 2, and those with a high level of perceived leadership should be assigned a score of 3.

3.3.2 SCORING & RANKING

Table 3-5 shows a summary of the scoring matrix used in HRDT. The value in parenthesis below each factor represents the weight associated with that factor, and was selected based on discussions with GDOT personnel, as well as the research team's judgment. It should also be noted that if the user interacting with HRDT has reason to override the estimates provided by HRDT, he/she can do so by manually selecting what is deemed to be a more appropriate score.

Once the scores are assigned, they are then multiplied with their respective weights, summed and scaled, and an overall score is then calculated for each of the tasks. Under this scheme, the maximum possible score is 12, and the minimum possible score is 4. Mathematically, the overall score for each task (S_t) can be expressed as follows:

$$S_t = n \times \sum_{i=1}^{n} W_i \times S_i$$

Where n is the number of factors (in this case, n=4), W_i is the weight, and S_i is the score for each factor. For example, consider the task of evaluating knowledge loss risk. Assume the employee has the following scores for each factor:

- Vacancy risk = 3 (projected retirement within 1 year)
- Uniqueness = 2 (moderately unique)
- Criticality = 3 (critical)
- Resource availability = 2 (some organizational support)

Table 3-5 – Scoring matrix for multivariate analysis

Task / Factor	KLR Assessment (1,3)	Mentor Evaluation (1)	Protégé Evaluation (1)	Position Evaluation (2,4)	Candidate Search / Experience Score (2,3)	Best Candidate (2,3)	Trainer Search (4)	Trainee Search (4)
1	Vacancy Risk (0.479)	***Willingness*** (0.500)	***Past Performance*** (0.521)	Absence Impact (0.479)	Position Level (0.521)	Experience Score (0.479)	Position Evaluation (0.250)	Position Level (0.400)
2	Uniqueness* (0.229)	KLR (0.250)	***Willingness*** (0.271)	Uniqueness (0.229)	Centrality (0.271)	***Willingness / Attitude*** (0.229)	Tenure at Current Position (0.250)	***Skill set*** (0.200)
3	Criticality* (0.229)	***Time Period*** (0.125)	***Time Availability*** (0.146)	Criticality (0.229)	Tenure at Current Position (0.146)	***Performance Score*** (0.229)	***Willingness / Attitude*** (0.250)	***Willingness / Attitude*** (0.200)
4	***Resource Availability*** (0.063)	***Time Availability*** (0.125)	***Location*** (0.063)	***Resource Availability*** (0.063)	***Certifications*** (0.063)	***Leadership*** (0.063)	***Time Availability*** (0.250)	***Time Availability*** (0.200)

(1) Used in Job Shadowing
(2) Used in Workforce Planning
(3) Used in Succession Planning
(4) Used in Cross Training

The factors in ***bold italics*** represent user-input values.

* For Job Shadowing, these values are user input, while for the Workforce Planning and Succession Planning, they are automatically calculated.

The overall KLR score (S_{KLR}) can then be calculated as:

$$S_{KLR} = 4 \times [(0.479 \times 3) + (0.229 \times 2) + (0.229 \times 3) + (0.063 \times 2)] = 10.8$$

Lastly, a "rating" system has been developed by studying all possible combinations of overall scores to define ranges, and applying additional constraints considering particular situations. For example, one constraint is that a candidate or trainee's overall score is automatically set to low if his/her willingness / attitude is chosen as low, as this indicates that he/she is not a willing participant and therefore not suitable for selection. Another constraint is that if both uniqueness and criticality are low, then the KLR score is automatically set to low. This rating system applies to KLR assessment, mentor and protégé evaluation, candidate search / experience score, candidate evaluation, and trainer and trainee evaluation. A summary of the overall scoring and rating criteria is provided in Table 3-6:

Table 3-6 – Overall scoring and rating criteria used in HRDT

Overall Score Range	Description
10 – 12	High Score
8 – 9.9	Moderate / Medium Score
4 – 7.9	Low Score

3.4 HRDT MODULES

There are a total of seven modules in HRDT as shown in Figure 3-1. A brief overview of each module is provided below. The User's Manual in Appendix A provides instructions on how to use each module.

3.4.1 JOB SHADOWING

The framework for the Job Shadowing module is identical to the one developed previously for the JOB SEEKER project, and is summarized in Figure 3-2.

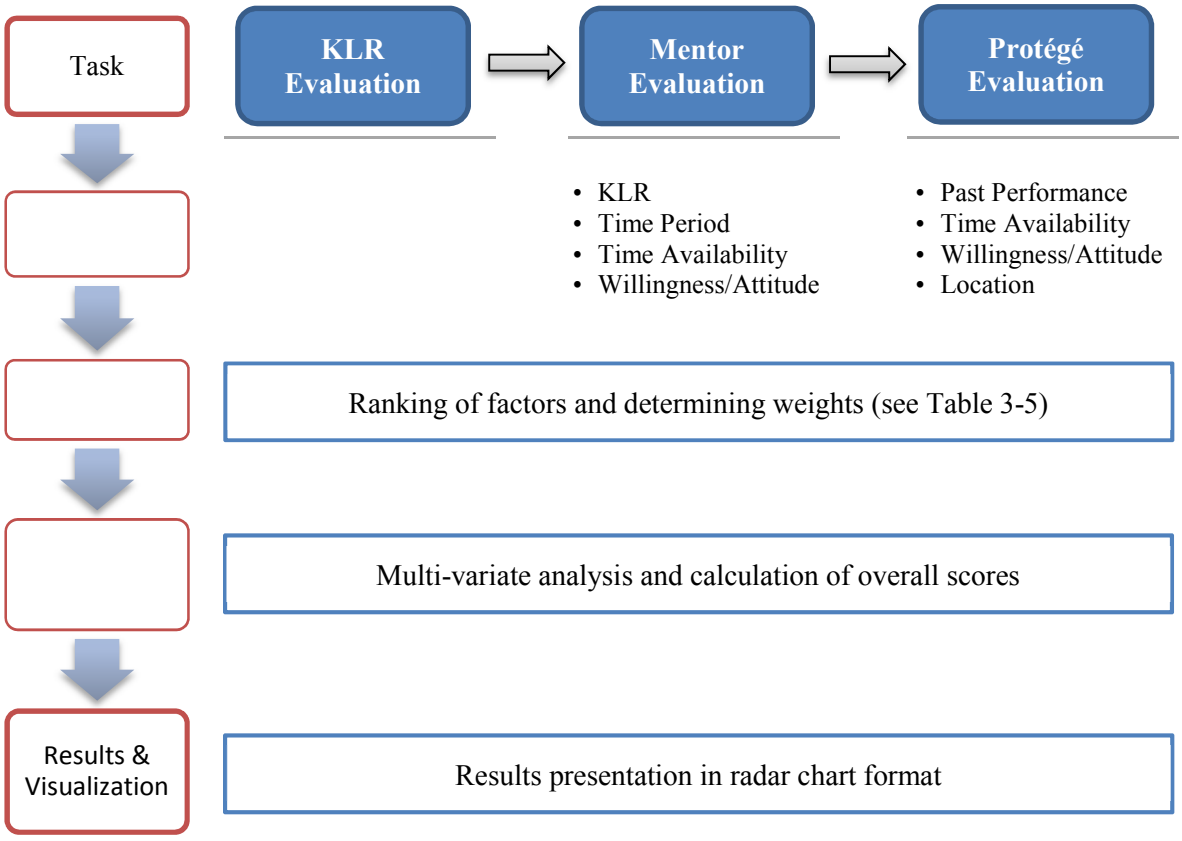


Figure 3-2 – Framework and evaluation process for the Job Shadowing module

3.4.2 WORKFORCE PLANNING

The framework for the Workforce Planning module is identical to Job Shadowing, while considering tasks and factors that are specific to this module, as summarized in Figure 3-3. For Workforce Planning, the focus is on identifying important positions (as predicted by the Position Evaluation task), as well as identifying potential and best candidates in case the important position must be replaced or supplemented.

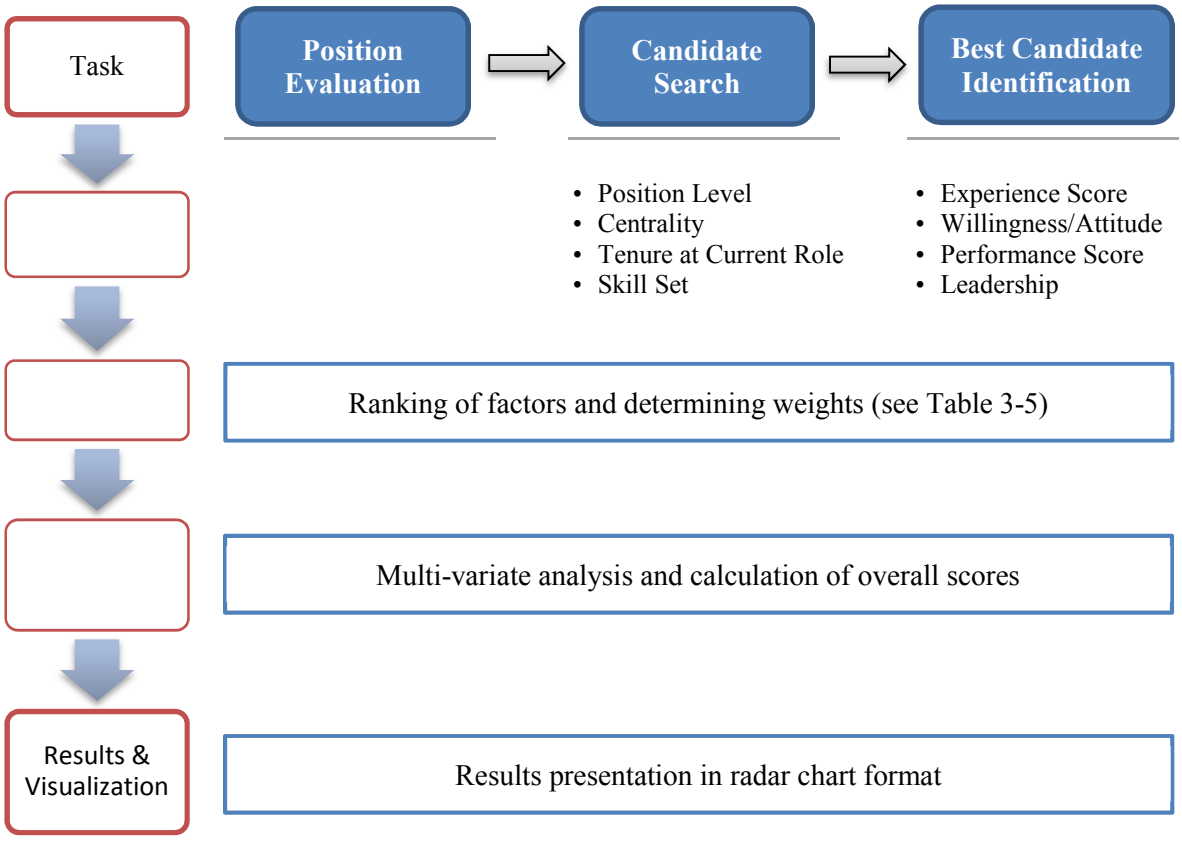


Figure 3-3 – Framework and evaluation process for the Workforce Planning module

3.4.3 SUCCESSION PLANNING

The framework for the Succession Planning module is identical to Workforce Planning, as summarized in Figure 3-4. For Succession Planning, the focus is on knowledge loss risk (as predicted by the KLR Evaluation task) instead of position importance, as well as identifying potential successor(s).

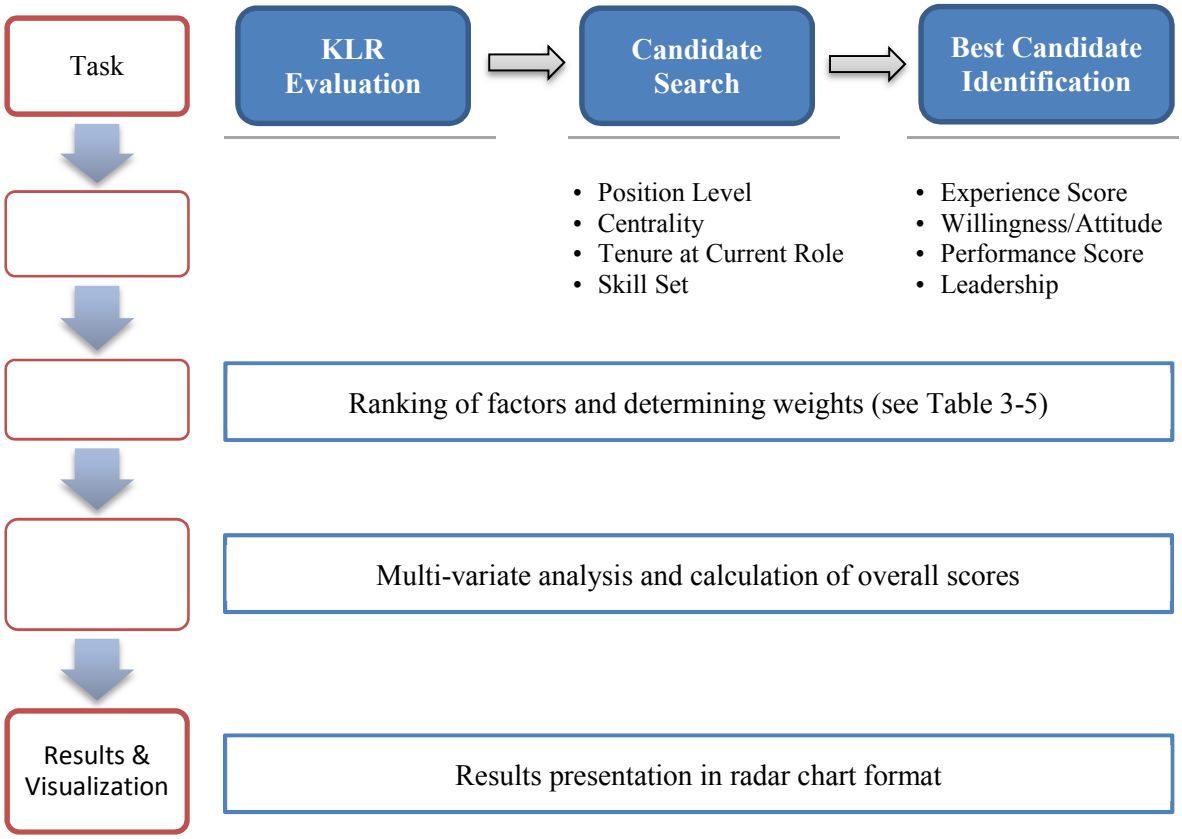


Figure 3-4 – Framework and evaluation process for the Succession Planning module

3.4.4 CROSS-TRAINING

The framework for the Cross Training module is identical to the previous modules, as summarized in Figure 3-5. The focus is on identifying potential positions, as well as trainer and trainees(s), to participate in cross-training.

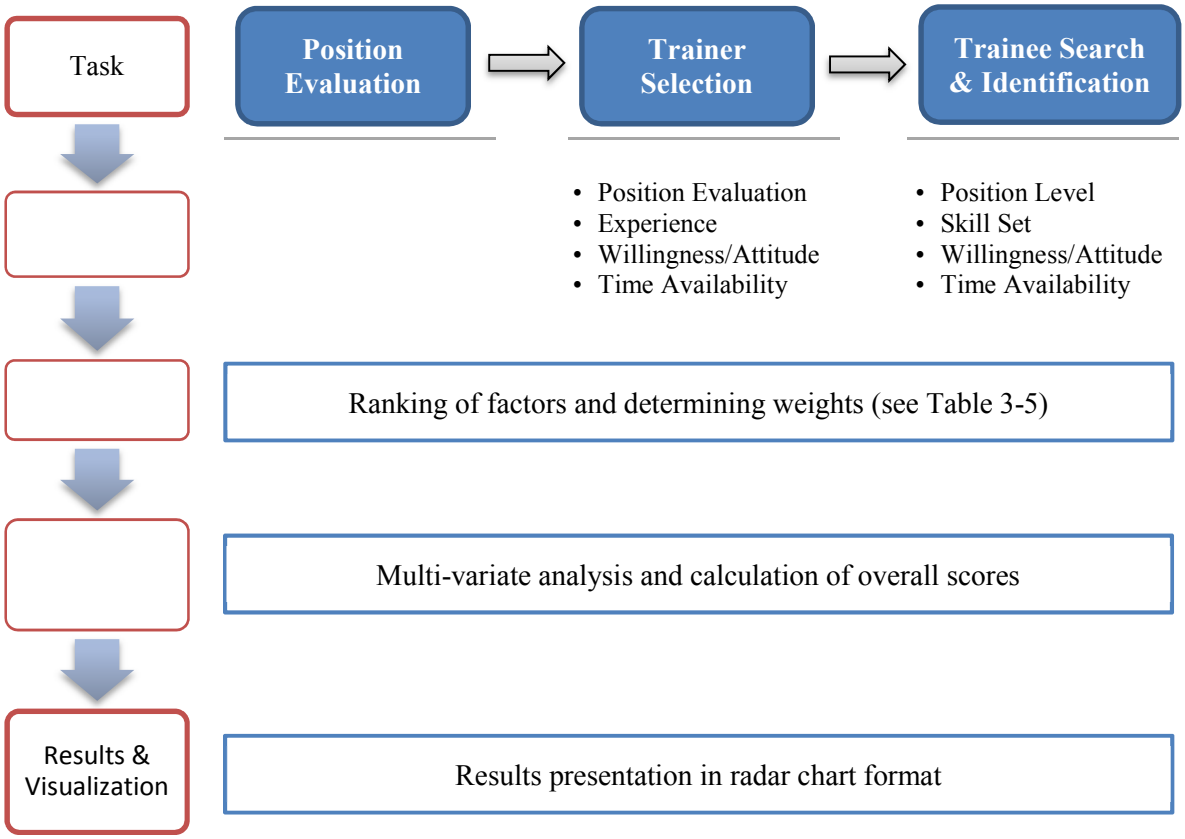


Figure 3-5 – Framework and evaluation process for the Cross Training module

3.4.5 TRAINING & DEVELOPMENT

A Training and Development (T&D) module was developed to allow GDOT personnel to evaluate different T&D alternatives based on various environmental characteristics as shown in Table 3-7. Default coefficients to quantify the effectiveness of each T&D method are provided in the table, and the users have the option to change these coefficients if desired. Based on the selections made, the module will provide suggestions for preferred T&D methods, ranked from most suitable (i.e. highest score) to the least suitable for the task (see Figure 3-6).

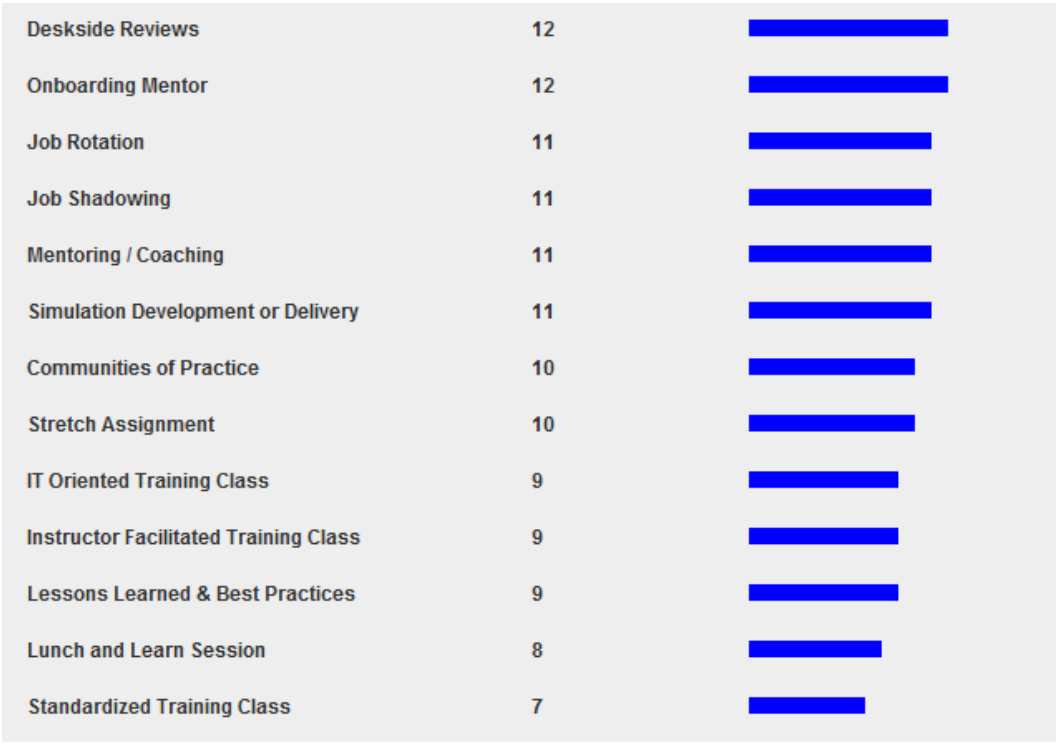


Figure 3-6 – Sample output from the Training & Development module

Table 3-7 – Environmental characters, T&D techniques and their effectiveness

	1	2	3	4	5	6	7	8	9	10	11	12	13
Environmental Characterization	Lessons Learned & Best Practices	Communities of Practice	Instructor Facilitated Training Class	Lunch and Learn Session	Standardized Training Class	Information Technology (IT) Oriented Training	Deskside Reviews	Job Shadowing	Simulation Development or Delivery	Job Rotation	Mentoring / Coaching	Stretch Assignment	Onboarding Mentor
Only one knowledge source and one knowledge receiver	2	2	2	1	1	2	3	3	2	3	3	3	3
Only one knowledge source and several to many knowledge receivers	2	2	3	2	1	2	2	2	2	3	2	2	2
Knowledge source is available less than 8 hours a week	1	1	1	1	1	1	2	1	1	1	1	1	3
Knowledge source is available between 8 and 16 hours a week	2	2	2	2	1	2	2	2	2	2	3	2	2
Knowledge source is available more than 16 hours a week	3	3	3	3	1	2	2	3	2	3	3	3	3
IT structure exists to support / distribute knowledge	2	3	2	2	1	2	2	2	2	2	2	2	3
IT structure does not exist to support / distribute knowledge	1	1	2	2	1	1	1	2	1	2	2	1	1
Knowledge source and knowledge receiver are co-located	2	2	3	2	2	2	3	3	3	3	3	3	2
Knowledge source and knowledge receiver are not co-located	1	1	2	1	2	2	2	1	1	2	2	2	1
There is less than 3 months available for knowledge transfer to take place	2	2	1	2	2	2	2	2	3	2	2	1	1
There is 3 to 6 months available for knowledge transfer to take place	2	2	2	1	2	2	2	2	1	2	2	2	2
There is more than 6 months available for knowledge transfer to take place	1	2	2	1	1	1	3	3	1	3	3	3	1

1 = Low effectiveness 2 = Moderate effectiveness 3 = High effectiveness

3.4.6 NETWORK ANALYSIS

The Network Analysis module allows for visualization of the network analyses performed for HRDT, including the various different centralities described in Section 3.2. It also allows users to view the back-constructed traditional organizational charts for the 58 predefined departmental units (based on the Master Organizational Chart). As an example, the organizational chart for the Office of Planning using the data extracted from the provided employee database is shown in Figure 3-7.

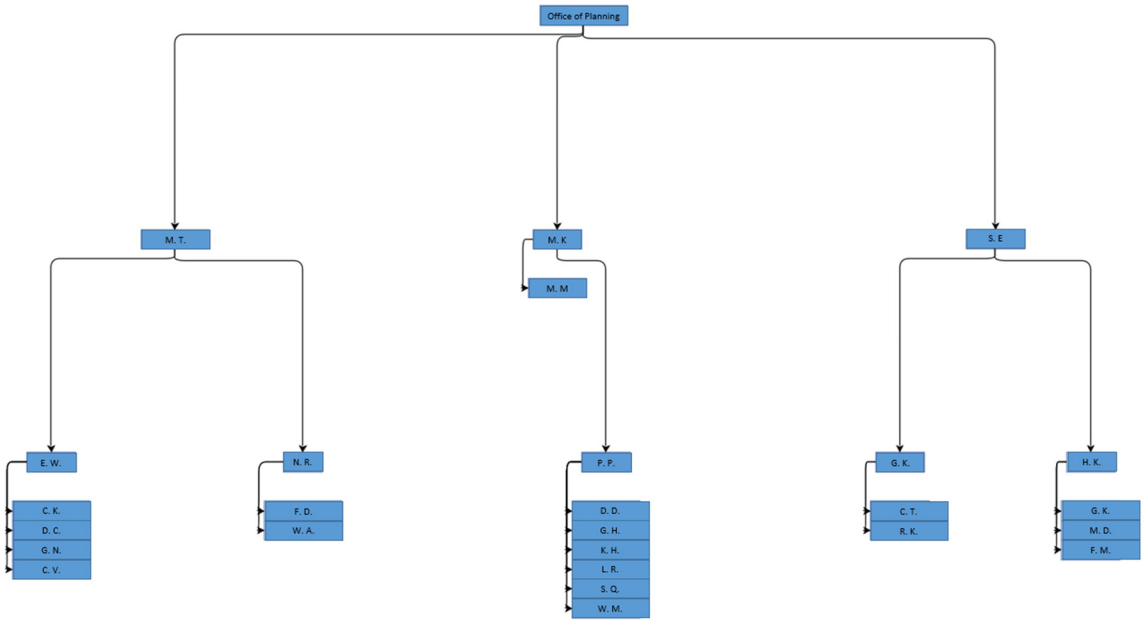


Figure 3-7 – An example of back-constructed traditional organizational chart from provided employee database

In addition, the user can view each departmental unit in a network representation, instead of the traditional organizational chart visualization. By representing the organizational chart as a network, it is possible not only to see who reports to whom, but also to visualize the connectivity of the individuals within the network. For example, the network representation of the organization chart shown above is given in Figure 3-8:

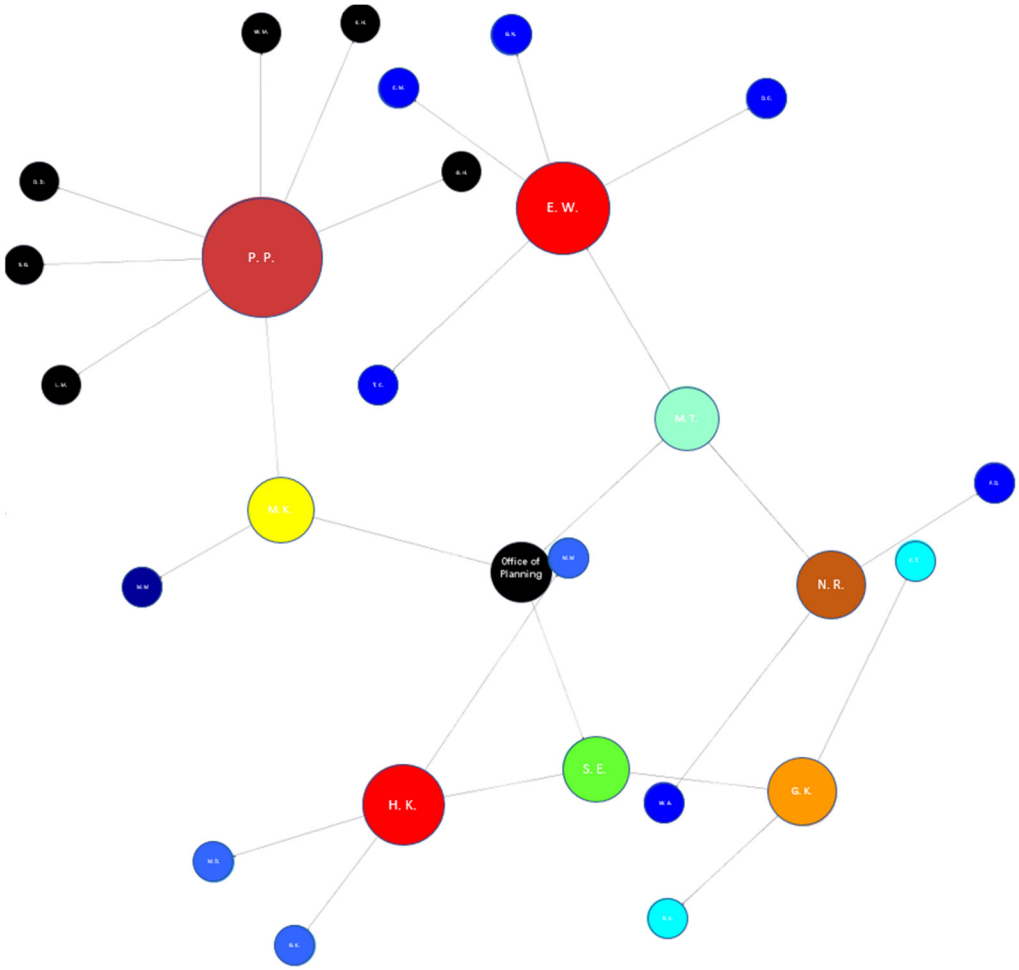


Figure 3-8 – Color-coded network representation of the organizational chart shown in Figure 3-7 (local PageRank centrality is shown, where larger circles indicate higher centrality and vice versa)

In this example, the size of the each circle is related to its centrality (larger the circle, higher the centrality). Clicking on each box in the organizational chart visualization or each circle in the network visualization displays the relevant attributes associated with that employee. More details are provided in the User's Manual in Appendix A.

3.4.7 DECISION MAKING

The Decision Making module in HRDT is an interactive dashboard which allows the operator to use spatial, tabular and graphical tools to visualize information rapidly to aid in decision making.

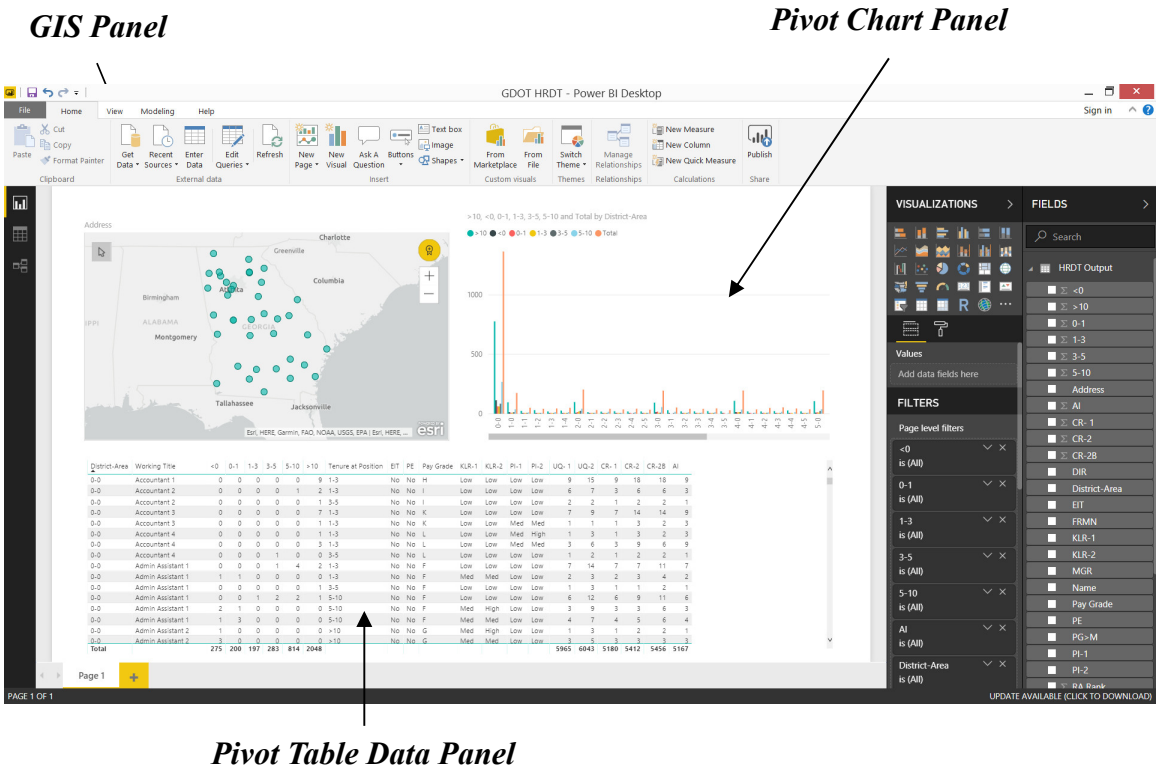


Figure 3-9 – Decision Making module overview

The Decision Making module has been integrated with Microsoft Power BI ® Desktop in order to leverage the pivot table / chart and spatial analysis and visualization capabilities of Power BI. An output file is generated by HRDT, which in turn acts as input into Power BI for subsequent analysis. The Decision Making module comprises three main panels: geographic information system panel – referred to as GIS panel (top left), pivot chart panel (top right), and pivot table data panel (bottom).

The Decision Making module provides powerful filtering and visualization tools, and allows the user to perform "what-if" scenario analyses in order to evaluate specific organizational needs. Table 3-8 provides a brief description of the parameters available for analysis using the Decision Making module.

Pivot charts/tables and spatial analysis tools in the Decision Making module are interlinked, meaning that performing map-based or chart-based analysis also updates the results shown in the other panels and vice versa. The User's Manual provides some practical examples demonstrating the decision-making applications of this module.

Table 3-8 – Summary of parameters available in the Decision Making module

Parameter	Description
Name	Employee Name
District-Area	The district and area to which the employee is assigned
Address	The physical address of the office to which the employee is assigned (based on district-area)
Working Title	The job title of the employee
Total	A value of "1" is assigned to each employee; acts as a counter for subsequent analyses
<0	A value of "1" is assigned to each employee who has already met the retirement requirements (based on age and/or tenure), but has not yet retired
0–1	A value of "1" is assigned to each employee who is expected to retire in the next one year (based on age and/or tenure)
1–3	A value of "1" is assigned to each employee who is expected to retire in the next one to three years (based on age and/or tenure)
3–5	A value of "1" is assigned to each employee who is expected to retire in the next three to five years (based on age and/or tenure)
5–10	A value of "1" is assigned to each employee who is expected to retire in the next five to ten years (based on age and/or tenure)
>10	A value of "1" is assigned to each employee who is expected to retire in more than ten years (based on age and/or tenure)
PE	"Yes" if the employee is a registered professional engineer
EIT	"Yes" if the employee is a registered engineer-in-training.
Pay Grade	The alphabetical pay grade of the employee as obtained from the database
Tenure at Position	The tenure of the employee (in years) at his/her most recent position, based on the position entry date in the database
VR	Vacancy risk of each employee as discussed in Report Section 3.2
RA	Resource availability (default value of 2)
UQ-1	Basic uniqueness of each employee as discussed in Report Section 3.2.1
UQ-2	Network-based uniqueness of each employee as discussed in Report Section 3.2.1

Table 3-8 (cont.) – Summary of parameters available in the Decision Making module

Parameter	Description
CR-1	Basic criticality of each employee as discussed in Report Section 3.2.2
CR-2	Network-based criticality of each employee using Local PageRank, as discussed in Report Section 3.2.2
CR-2B	Network-based criticality of each employee using Cosine Dissimilarity, as discussed in Report Section 3.2.2
AI	Absence impact of each employee as discussed in Report Section 3.2.3
KLR-1	Knowledge loss risk of each employee, calculated using the Basic definitions of uniqueness and criticality
KLR-2	Knowledge loss risk of each employee, calculated using the Network-based definitions of uniqueness and criticality
FRMN	"True" if an employee's job title contains FRMN (Foreman); "False" otherwise
SPI	"True" if an employee's job title contains SPI (Superintendent); "False" otherwise
Spv	"True" if an employee's job title contains Spv (Supervisor); "False" otherwise
MGR	"True" if an employee's job title contains MGR (Manager); "False" otherwise
DIR	"True" if an employee's job title contains DIR (Director); "False" otherwise
Tenure => 2 yrs	"True" if an employee has been at his/her most recent position for at least two years

4. CONCLUSIONS

This report has presented the concepts and framework behind Human Resources Data Tool (HRDT), a Java-based software package developed specifically for GDOT to aid in analysis of HR data and for use in planning and decision making. A User's Manual for the software has also been prepared and is attached as Appendix A. A Training Module presentation has also been prepared and is attached as Appendix B. This project builds upon the previous JOB SEEKER project by applying data mining and data analysis techniques such as network analysis, spatial analysis, and pivot charts and tables, in order to gain a deeper understanding and insight into the existing GDOT HR data.

HRDT includes a modular and consistent framework for evaluating various parameters which are important to planning and decision making. The software includes seven data modules:

- Job Shadowing
- Workforce Planning
- Succession Planning
- Cross Training
- Training and Development
- Network Analysis
- Decision Making

These data modules can allow GDOT HR personnel to assess workforce and training needs not only at an individual level, but they can also be used for planning and strategic decision-making at the organizational level. It is anticipated that HRDT will help GDOT advance from being an information-rich organization to a knowledge-rich one, in turn helping GDOT achieve its stated mission of delivering "a transportation system focused on innovation, safety, sustainability and mobility" and the associated stated goal to "recruit, train and retain a quality workforce".

5. REFERENCES

Barthelemy, M., 2004. Betweenness centrality in large complex networks. *European Physics Journal B*, Volume 38, pp. 163-168.

Brin, S. and Page, L., 1998. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30: 107-117.

Butts, C. T., 2008. Social network analysis: A methodological introduction. *Asian Journal of Social Psychology*, Volume 11, pp. 13-41.

Freeman, L. C., 1978/1979. Centrality in Social Networks: Conceptual Clarification. *Social Networks*, pp. 215-239.

Frost, J.D., 2016. "JOB SEEKER" (Job Shadowing for Employee Engagement through Knowledge and Experience Retention). Final Report, GDOT Research Project No. RP13-12, Report No. FHWA-GA-16-1312.

Han, J., Kamber, M. and Pei, J., 2012. Data mining: Concepts and Techniques. Third Edition, New York: Morgan-Kaufman.

Hanneman, R.A. and Riddle. M., 2005. Introduction to social network methods. Riverside, CA: University of California, Riverside (published in digital form at http://faculty.ucr.edu/~hanneman/)

Hansen, M. T., 1999. The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits. *Administrative Science Quarterly*, 44, pp. 82-111.

Hansen, M. T., Mors, M. L. & Lovas, B., 2005. Knowledge Sharing in Organizations: Multiple Networks, Multiple Phases. *The Academy of Management Journal*, 48(5), pp. 776-793.

Jantan, H., Othman, Z.A. and Hamdan, A.R., 2010. Human Talent Prediction in HRM using C4.5 Classification Algorithm. *International Journal of Advanced Trends in Computer Science and Engineering*, 2(8): 2526-2534.

Ranjan, J., Goyal, D.P. & Ahson, S.I., 2008. Data mining techniques for better decisions in human resource management systems. *International Journal of Business Information Systems*, 3(5): 464-481.

Reagans, R. & McEvily, B., 2003. Network Structure and Knowledge Transfer: The Effects of Cohesion and Range. *Administrative Science Quarterly*, 48, pp. 240-267.

Sadath, L., 2013. Data Mining: A Tool for Knowledge Management in Human Resource. *International Journal of Innovative Technology and Exploring Engineering*, 2(6): 154-159.

Silwattananusarn, T. and Tuamsuk, K., 2012. Data Mining and Its Applications for Knowledge Management: A Literature Review from 2007 to 2012. *International Journal of Data Mining & Knowledge Management Process*, 2(5): 13-24.

APPENDIX A –

GDOT HR Data Tool User's Manual

User's Manual

March 2019

1. INTRODUCTION & OVERVIEW

1.1 INTRODUCTION

HR Data Tool (HRDT) is a Java-based program which performs network, spatial and multivariate analyses to aid GDOT HR personnel in planning and decision making. This User's Manual provides an overview of the program and describes how the various modules can be used.

1.2 INSTALLATION

The program can be installed using the provided installer package. Prior to installation, the user should make sure that his/her computer has the latest version of Java (minimum Java 8) installed. The installer will also install Microsoft Power BI Desktop onto the computer. This free software is required for the Decision Making module of HRDT.

1.3 OVERVIEW

Upon starting up HRDT, the following landing page is displayed, showing the seven modules associated with the program (Figure A-1):

- Decision Making
- Network Analysis
- Job Shadowing
- Workforce Planning
- Succession Planning
- Cross Training
- Training & Development

Figure A-1 – Application Landing Page

The modules shown on the right in Figure A-1 (Job Shadowing, Workforce Planning, Succession Planning, Cross Training, and Training & Development) can be broadly categorized as Multivariate Analyses. The other two main modules are Decision Making and Network Analysis. Multivariate Analyses are discussed in Section 2 of Appendix A, Network Analysis is discussed in Section 3 of Appendix A, and Decision Making is discussed in Section 4 of Appendix A.

It can also be seen from Figure A-1 that the users have the ability to update the database, which in turn will re-run the calculations based on the latest data. This button is to be used when there are changes in the GDOT employee database, for example new employees hired, existing employees left, employees promoted, etc. This process updates any intermediate files that are essential for the all parts of this application to function correctly.

Most importantly, users should update the database using the same Structured Query Language (SQL) statement as given in Table 3-3 of the main report, and strictly following the format in the current database, including the addition of the "PE" and "EIT" columns manually as column AL and AM, respectively. Any deviations from the current format will result in HRDT not functioning as intended.

2. MULTIVARIATE ANALYSIS MODULES

The Multivariate Analysis consists of five modules: Job Shadowing, Workforce Planning, Succession Planning, Cross Training and Training & Development. Each module has its own unique functionality, as discussed in the following sections.

2.1 JOB SHADOWING

The Job shadowing module helps HR personnel evaluate an employee's knowledge loss risk, referred to as KLR, as well determine the most suitable mentor and protégé for participation in a job shadowing program. Figure A-2 shows the user interface (UI) of the Job Shadowing Module.

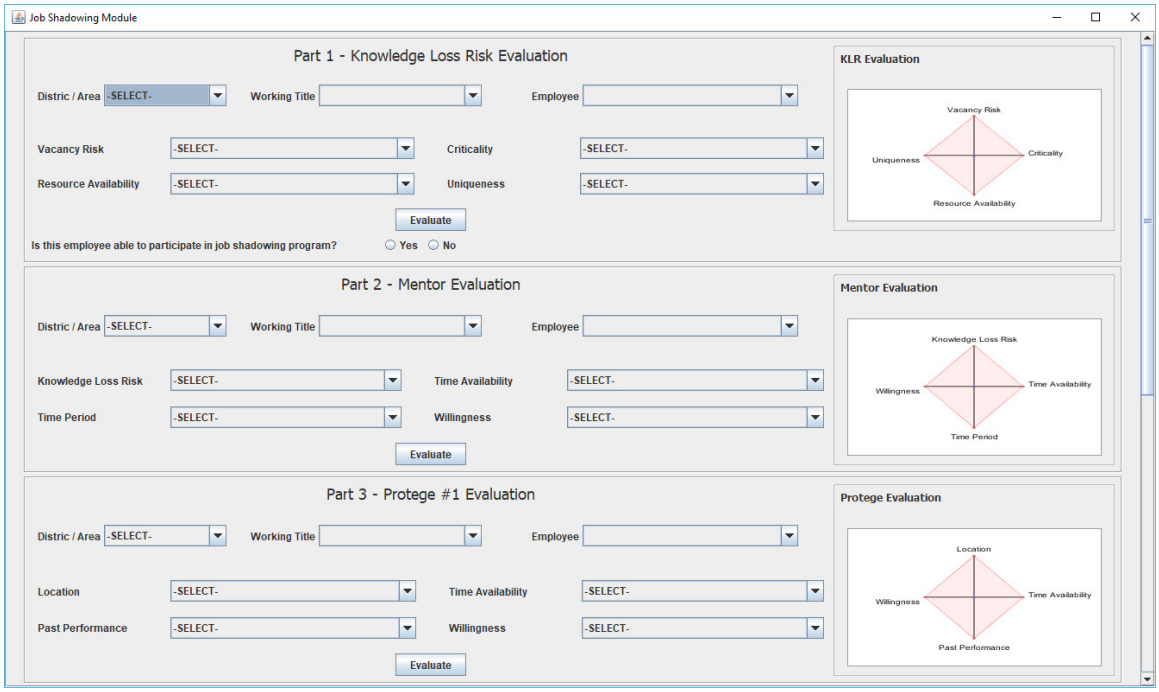


Figure A-2 – Job Shadowing UI

In Part 1, Knowledge Loss Risk Evaluation, users start filtering employees by selecting District / Area, then Working Title. When a specific employee is then selected, his/her vacancy risk is automatically populated based on information extracted from the database. This may be overridden manually by users if needed. Next, users assign the employee's Criticality, Uniqueness and Resource Availability from the dropdown lists. Once the four parameters are selected, users can click on the Evaluate button to evaluate the employee's KLR. The results are shown on the right with KLR score and radar chart visualization.

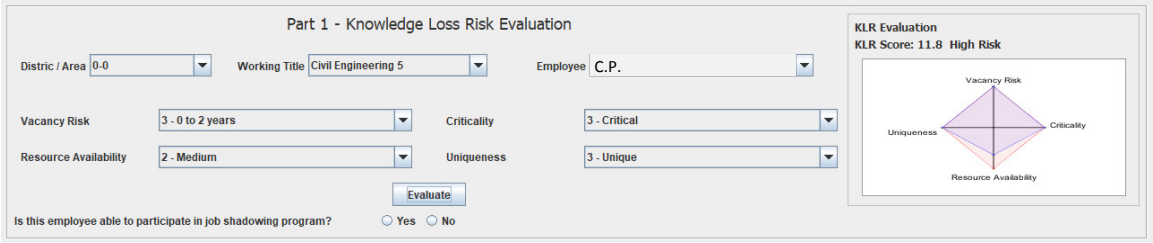


Figure A-3 – KLR Evaluation for Job Shadowing

If the employee (presumably with high KLR) is able to participate in the job shadowing program, clicking the "Yes" radio button will automatically import his/her basic information (District / Area, Working Title and name) and his/her KLR score into Part 2, Mentor Evaluation. If the "No" radio button is selected, then the mentor and his/her KLR would have to be manually selected in Part 2. However, it is anticipated that the vast majority of the time, the employee with a high KLR score (i.e., high attrition risk) will act as the mentor.

Next, users assign the mentor's Time Period, Time Availability and Willingness from the dropdown lists, and click the Evaluate button. The mentor evaluation results including the mentor score and the radar chart visualization are then shown on the right.

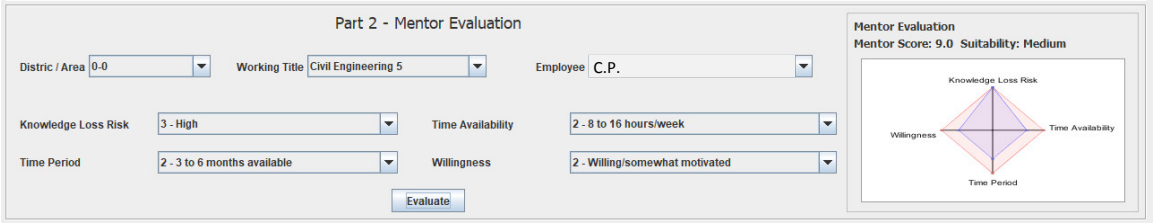


Figure A-4 – Mentor Evaluation for Job Shadowing

In part 3, Protégé Evaluation, users can evaluate up to three candidates. Users select candidate protégé in each division with the filtering function, then input his/her Location, Past Performance, Time Availability and Willingness to evaluate his/her suitability for being a protégé. After clicking the Evaluate button, the protégé evaluation results including the protégé score and the radar chart visualization are shown on the right.

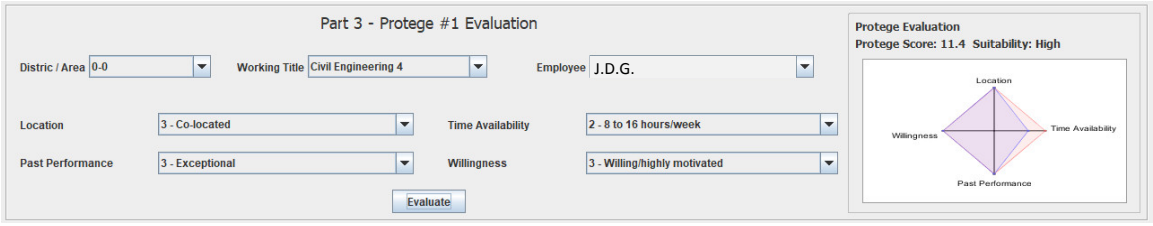


Figure A-5 – Protégé Evaluation for Job Shadowing

Lastly, the Generate Report button at the end of the module lets users print the entire assessment results (as a JPG or PNG image file) for future reference and/or for filing.

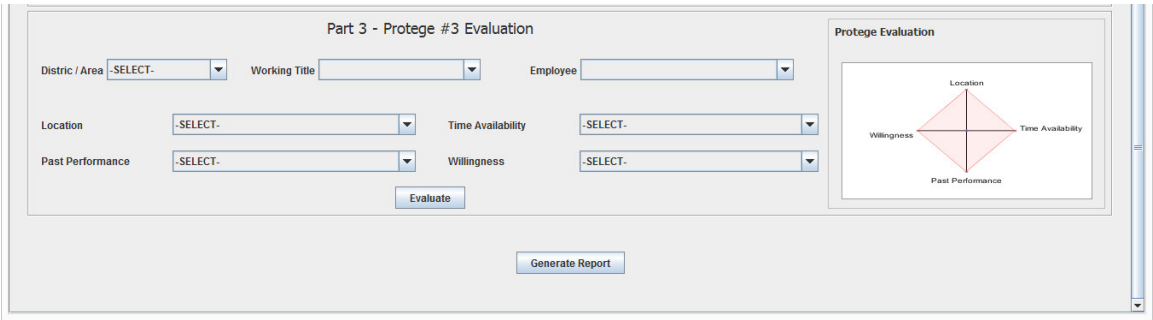


Figure A-6 – Generate Report Function for Job Shadowing

2.2 WORKFORCE PLANNING

Workforce Planning module helps HR personnel identify high-importance positions and the employees at those positions, and helps to evaluate suitable candidates for a selected position for better workforce planning. Figure A-7 shows the UI for the Workforce Planning Module.

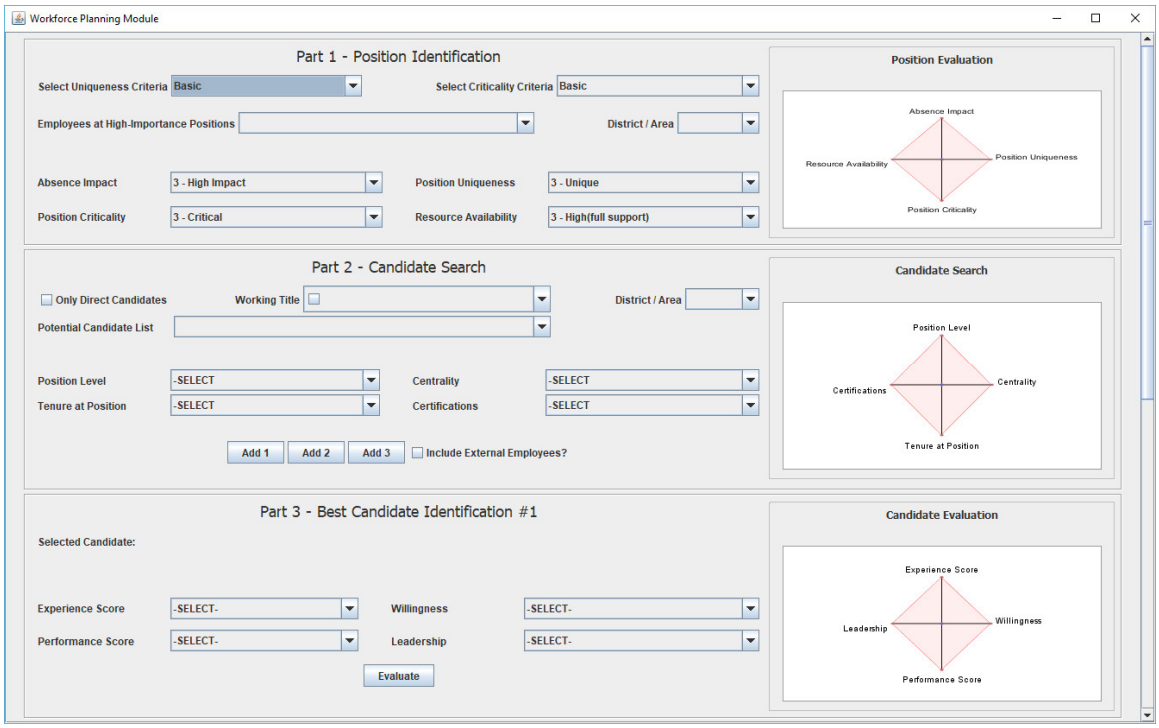


Figure A-7 – Workforce Planning UI

Part 1, Position Identification, aims to help users identify high-importance positions. In this part, users start by choosing the desired criteria, Basic or Network, for evaluating uniqueness and criticality of job positions. After the desired criteria are chosen, the employees at high-importance positions are returned, and the list of those employees are

shown in the dropdown list in alphabetical order of their last names. If desired, users may use the District / Area filter to further narrow down the list of employees based on their location.

Once an employee is selected from the dropdown list, his/her basic information is displayed right beneath, and the values of Absence Impact, Position Criticality and Position Uniqueness are automatically calculated and populated into their respective dropdown lists. The Resource Availability is set as a default value of 2 – Medium, which may be overridden manually. Also note that if the user does not agree with the suggested scores for the other three factors, they may also be overridden manually. The Position Evaluation results, including the calculated Position Score and radar chart visualization, are automatically shown on the right, and any changes in the four parameters will be reflected immediately in the evaluation results.

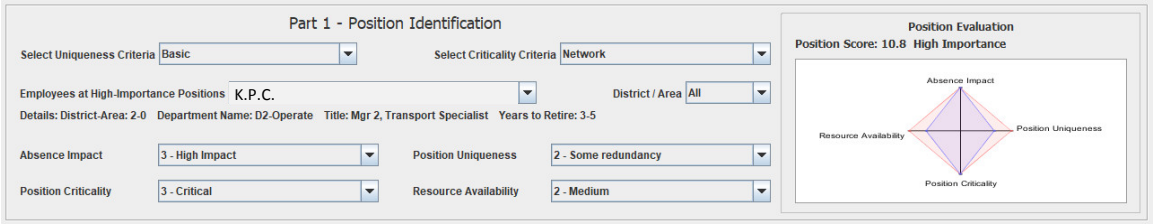


Figure A-8 – Position Identification for Workforce Planning

Part 2, Candidate Search, helps users search for potential candidates for the purpose of workforce planning with respect to the position identified in Part 1. In this part, selecting the "Only Direct Candidates" checkbox shows only the employees who report directly to the position identified in Part 1. By unchecking this box, the entire workforce can be searched for candidates via the Working Title and District / Area filters.

Once a candidate is selected from the dropdown list, his/her basic information is displayed right beneath, and the values of Position Level, Tenure at Position and Centrality are automatically calculated and populated into their respective dropdown lists. The Certifications is set as a default value of 2 – Medium, which may be overridden manually. Also note that if the user does not agree with the suggested scores for the other three factors, they may also be overridden manually. The Candidate Search results, including the calculated Experience Score and radar chart visualization, are automatically shown on the right, and any changes in the four parameters will be reflected immediately in the evaluation results.

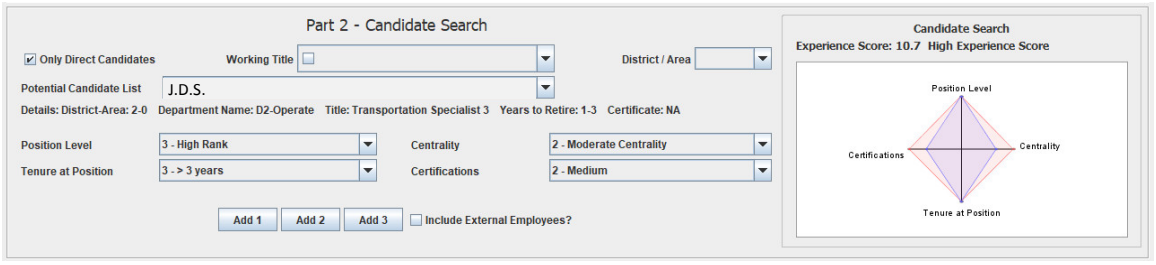


Figure A-9 – Candidate Search for Workforce Planning

Part 3, Best Candidate Identification, allows up to 3 potential candidates identified in Part 2 to be added to Part 3, using the Add 1/2/3 buttons. Once a potential candidate is added, his/her name and other relevant information are automatically populated, and his/her Experience Score is also populated into the first dropdown list automatically. Users then manually assign the Performance Score, Willingness, and Leadership, and click the Evaluate button. The Candidate Evaluation results, including the Candidate Score, his/her suitability and radar chart visualization are then shown on the right.

Figure A-10 – Best Candidate Identification for Workforce Planning

There is also the flexibility to add up to three external employees to the evaluation in Part 3 by checking the Include External Employees checkbox in Part 2, and entering the desired number of external employees to include. Additional evaluation fields will be added at the bottom of the page. Because these external candidates are not existing employees, and therefore their information is not contained in the database, users must manually input external employees' information to evaluate their suitability.

Figure A-11 – External Candidates for Workforce Planning

Lastly, the Generate Report button at the end of the module lets users print the entire assessment results (as a JPG or PNG image file) for future reference and/or for filing.

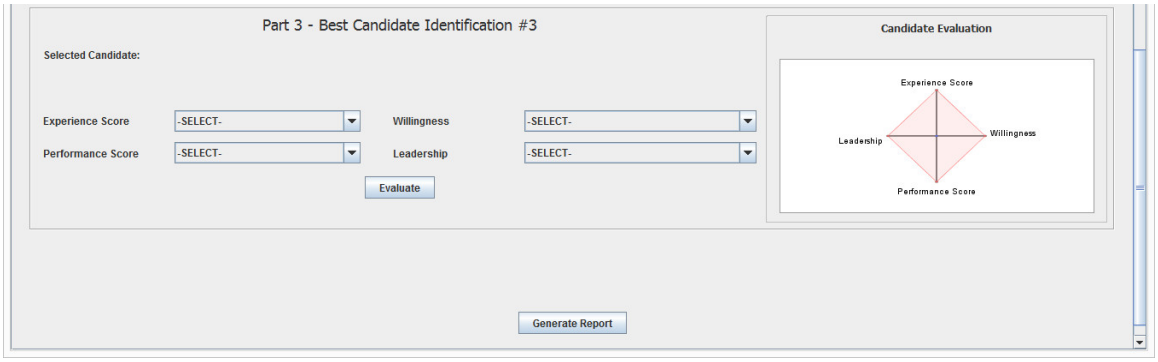


Figure A-12 – Generate Report Function for Workforce Planning

2.3 SUCCESSION PLANNING

Succession Planning module helps HR personnel identify employees with high knowledge loss risk (referred as KLR), and helps to evaluate suitable candidates for a selected position for succession planning purposes. Figure A-13 shows the UI for the Succession Planning Module.

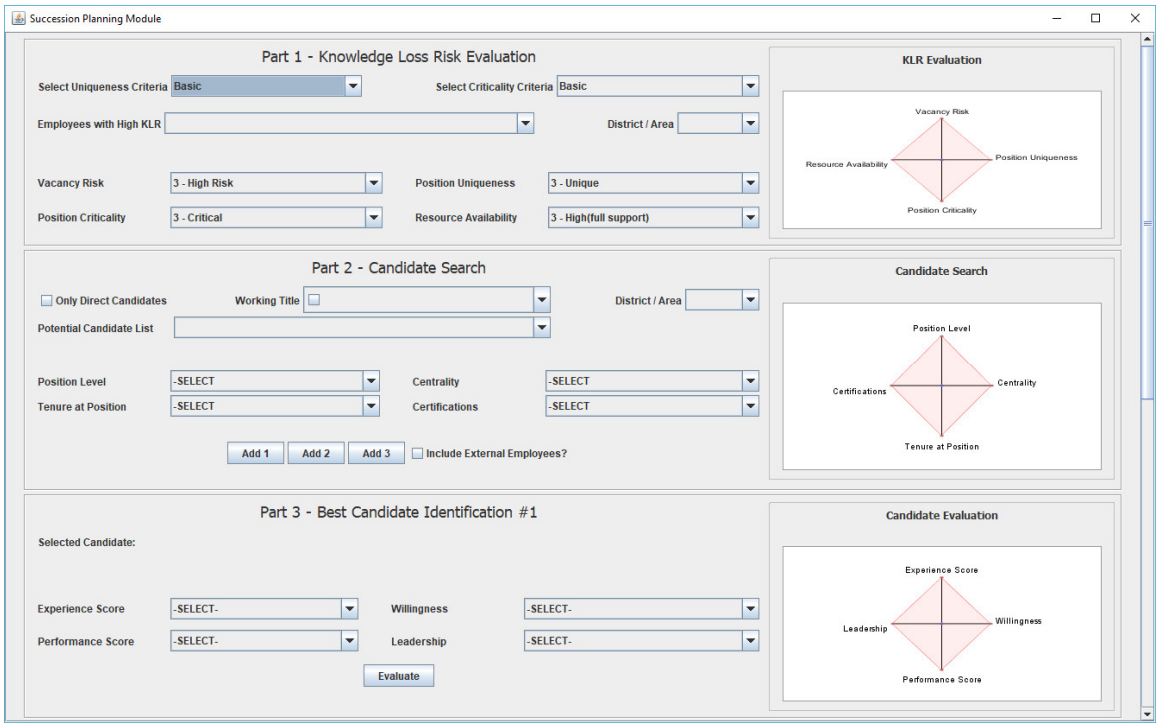


Figure A-13 – Succession Planning UI

Part 1 aims to help users identify employees with high KLR. In this part, users start by choosing the desired criteria, Basic or Network, for evaluating uniqueness and criticality of the employees. After the desired criteria are chosen, the employees with high KLR are returned, and the list of those employees are shown in the dropdown list in

alphabetical order of their last names. If desired, users may use the District / Area filter to further narrow down the list of employees based on their location.

Once an employee is selected from the dropdown list, his/her basic information is displayed right beneath, and the values of Vacancy Risk, Position Criticality and Position Uniqueness are automatically calculated and populated into their respective dropdown lists. The Resource Availability is set as a default value of 2 – Medium, which may be overridden manually. Also note that if the user does not agree with the suggested scores for the other three factors, they may also be overridden manually. The KLR evaluation results, including the calculated KLR Score and radar chart visualization, are automatically shown on the right, and any changes in the four parameters will be reflected immediately in the evaluation results.

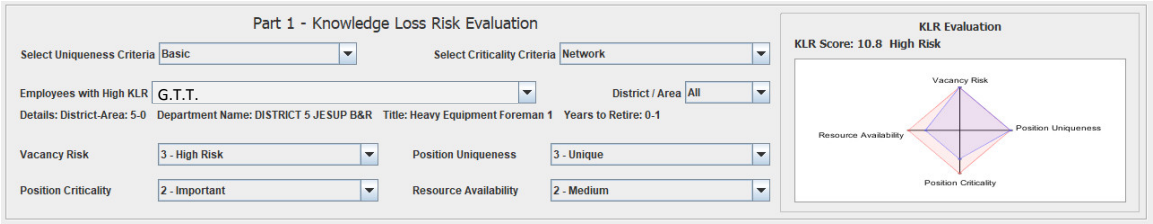


Figure A-14 – KLR Evaluation for Succession Planning

Part 2, Candidate Search, helps users search for potential candidates for the purpose of succession planning with respect to the position identified in Part 1. In this part, selecting the "Only Direct Candidates" checkbox shows only the employees who report directly to the person identified in Part 1. By unchecking this box, the entire workforce can be searched for candidates via the Working Title and District / Area filters.

Once a candidate is selected from the dropdown list, his/her basic information is displayed right beneath, and the values of Position Level, Tenure at Position and Centrality are automatically calculated and populated into their respective dropdown lists. The Certifications is set as a default value of 2 – Medium, which may be overridden manually. Also note that if the user does not agree with the suggested scores for the other three factors, they may also be overridden manually. The Candidate Search results, including the calculated Experience Score and radar chart visualization, are automatically shown on the right, and any changes in the four parameters will be reflected immediately in the evaluation results.

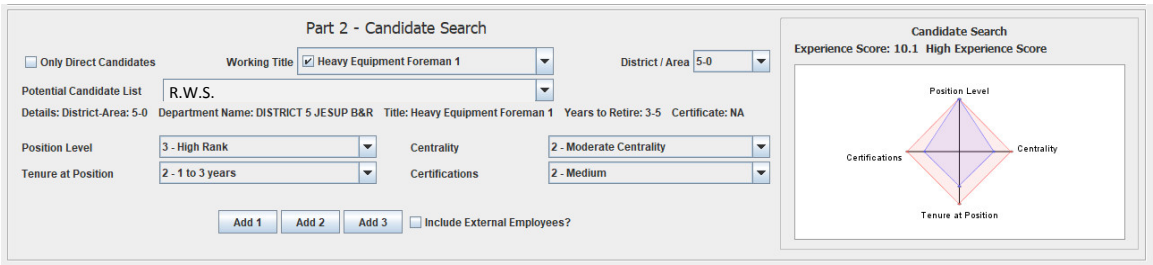


Figure A-15 – Candidate Search for Succession Planning

Part 3, Best Candidate Identification, allows up to 3 potential candidates identified in Part 2 to be added to Part 3, using the Add 1/2/3 buttons. Once a potential candidate is added, his/her name and other relevant information are automatically populated, and his/her Experience Score is also populated into the first dropdown list automatically. Users then manually assign the Performance Score, Willingness, and Leadership, and click the Evaluate button. The Candidate Evaluation results, including the Candidate Score, his/her suitability and radar chart visualization, are then shown on the right.

Figure A-16 – Best Candidate Identification for Succession Planning

There is also the flexibility to add up to three external employees to the evaluation in Part 3 by checking the Include External Employees checkbox in Part 2, and entering the desired number of external employees to include. Additional evaluation fields will be added at the bottom of the page. Because these external candidates are not existing employees, and therefore their information is not contained in the database, users must manually input external employees' information to evaluate their suitability.

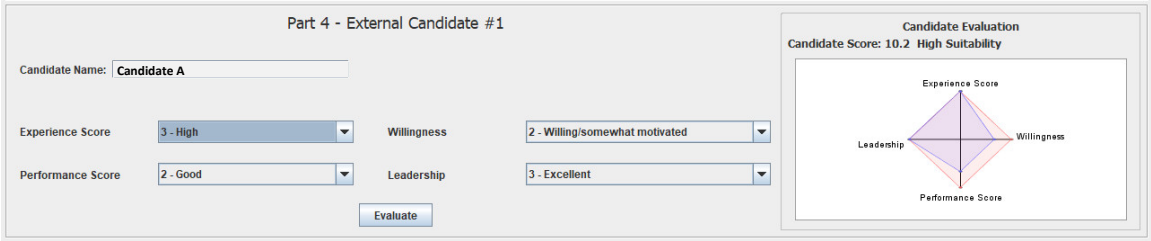


Figure A-17 – External Candidates for Succession Planning

Lastly, the Generate Report button at the end of the module lets users print the entire assessment results (as a JPG or PNG image file) for future reference and/or for filing.

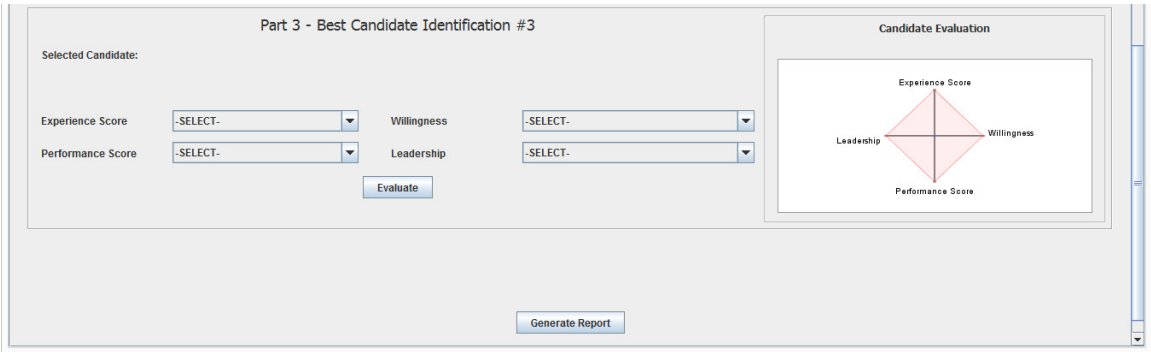


Figure A-18 – Generate Report Function for Succession Planning

2.4 CROSS TRAINING

Cross Training module helps HR personnel identify high-importance positions and the employees at those positions, and helps to evaluate suitable trainer and trainee candidate for cross training. Figure A-19 shows the user interface of the Cross Training Module.

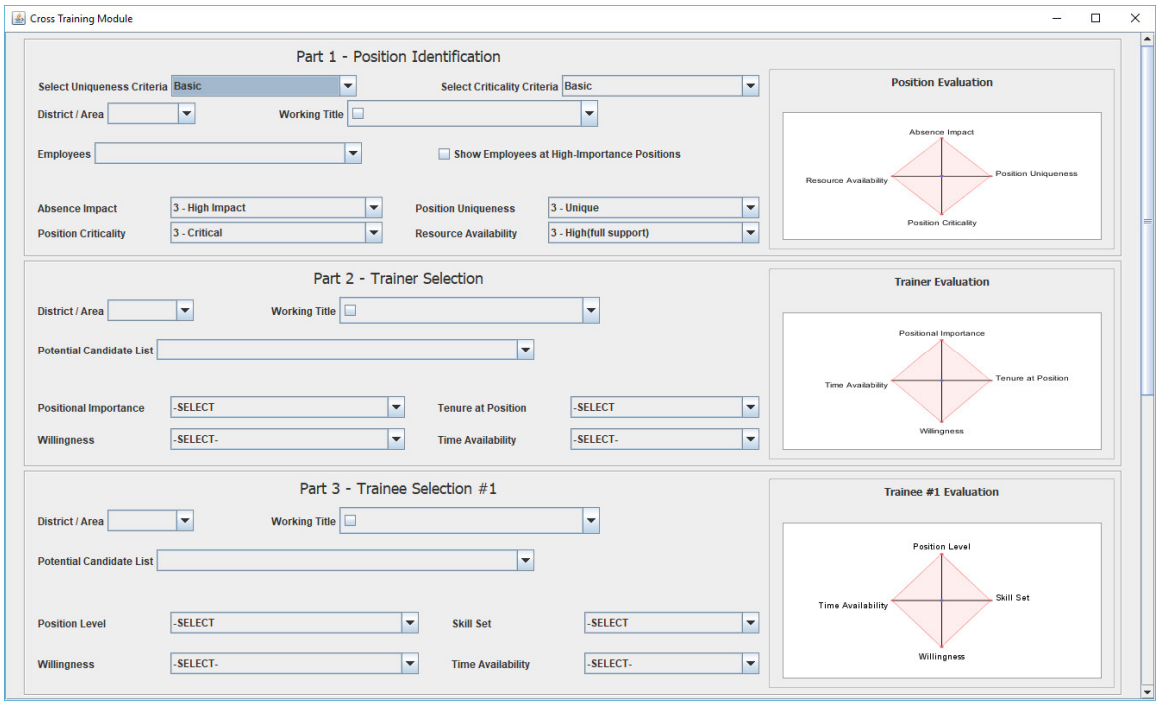


Figure A-19 – Cross Training UI

Part 1, Position Identification, aims to help users to quantify positional importance. In this part, users start by choosing the desired criteria, Basic or Network, for evaluating uniqueness and criticality of job positions. After the desired criteria are chosen, list of the employees are shown in the dropdown list in alphabetical order of their last names. If desired, users may use the District / Area and Working Title filters to further narrow

down the list of employees based on their location. Users also have the flexibility to show only the employees at high-importance positions in the dropdown by using the checkbox.

Once an employee is selected from the dropdown list, his/her basic information is displayed right beneath, and the values of Absence Impact, Position Criticality and Position Uniqueness are automatically calculated and populated into their respective dropdown lists. The Resource Availability is set as a default value of 2 – Medium, which may be overridden manually. Also note that if the user does not agree with the suggested scores for the other three factors, they may also be overridden manually. The Position Evaluation results, including the calculated Position Score and radar chart visualization, are automatically shown on the right, and any changes in the four parameters will be reflected immediately in the evaluation results.

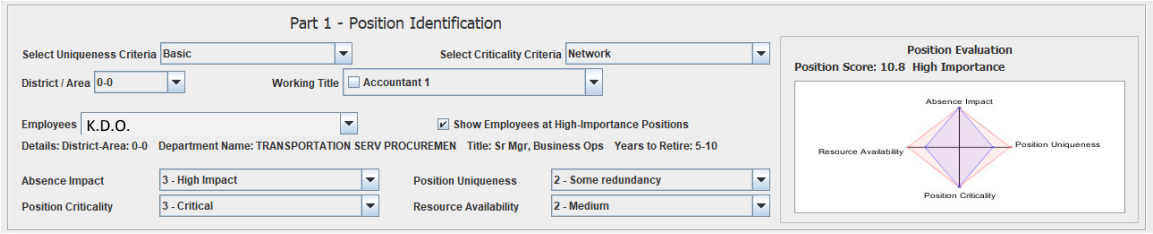


Figure A-20 – Position Identification for Cross Training

Part 2, Trainer Selection, helps users search for a potential trainer for the purpose of cross training with respect to the position identified in Part 1. Users may use the District / Area and Working Title filters to narrow down the employees in the returned potential candidate list.

Once a trainer candidate is selected from the dropdown list, his/her basic information is displayed right beneath, and the values of Position Level and Tenure at Position are automatically calculated and populated into their respective dropdown lists. The Willingness and Time Availability are set as a default value of 2 – Medium, which may be overridden manually. The Trainer Evaluation results, including the Trainer Score and radar chart visualization, are automatically shown on the right, and any changes in the four parameters will be reflected immediately in the evaluation results.

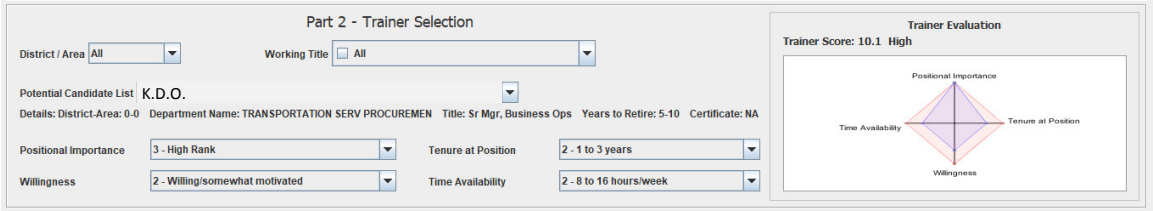


Figure A-21 – Trainer Selection for Cross Training

Part 3, Trainee Selection, helps users search for up to three potential trainees for the purpose of cross training with respect to the position identified in Part 1 and the trainer selected in Part 2. Users may use the District / Area and Working Title filters to narrow down the employees in the returned potential candidate list.

Once a trainee candidate is selected from the dropdown list, his/her basic information are displayed right beneath, and the values of Position Level are automatically calculated and input into the dropdown list. The Skill Set, Willingness and Time Availability are set as a default value of 2 – Medium, which may be overridden manually. The Trainee

Evaluation results, including the Trainee Score and radar chart visualization, are automatically shown on the right, and any changes in the four parameters will be reflected immediately in the evaluation results.

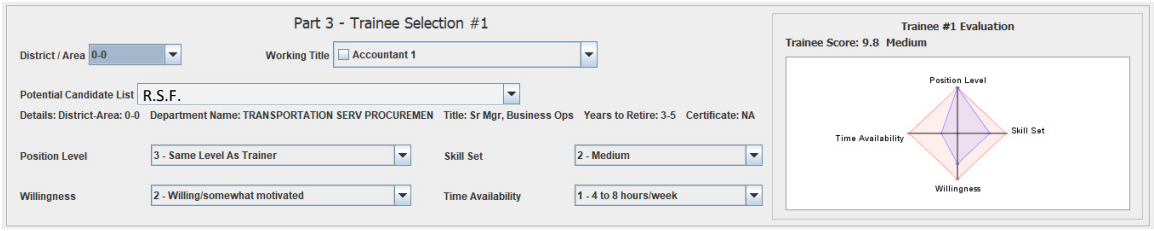


Figure A-22 – Trainee Selection for Cross Training

Lastly, the Generate Report button at the end of the module lets users print the entire assessment results (as a JPG or PNG image file) for future reference and/or for filing.

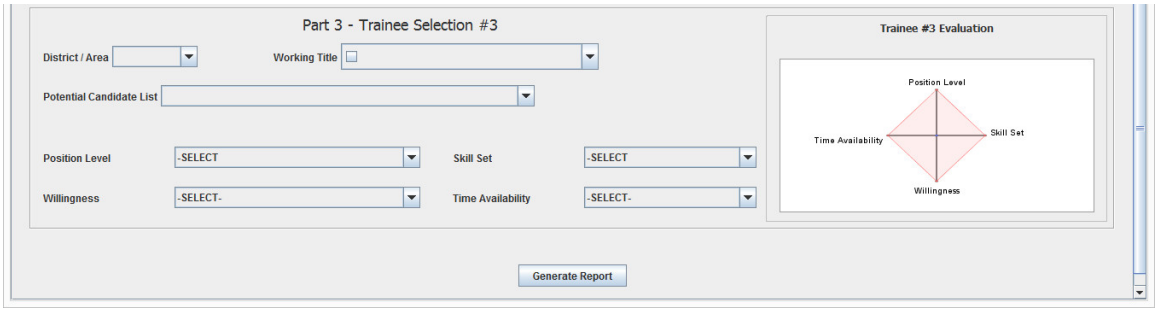


Figure A-23 – Generate Report Function for Cross Training

2.5 TRAINING & DEVELOPMENT

The Training & Development module helps HR personnel to evaluate the suitability of various training and development strategies while considering different environmental characteristics, through five simple questions in a simple UI, as shown in Figure A-24:

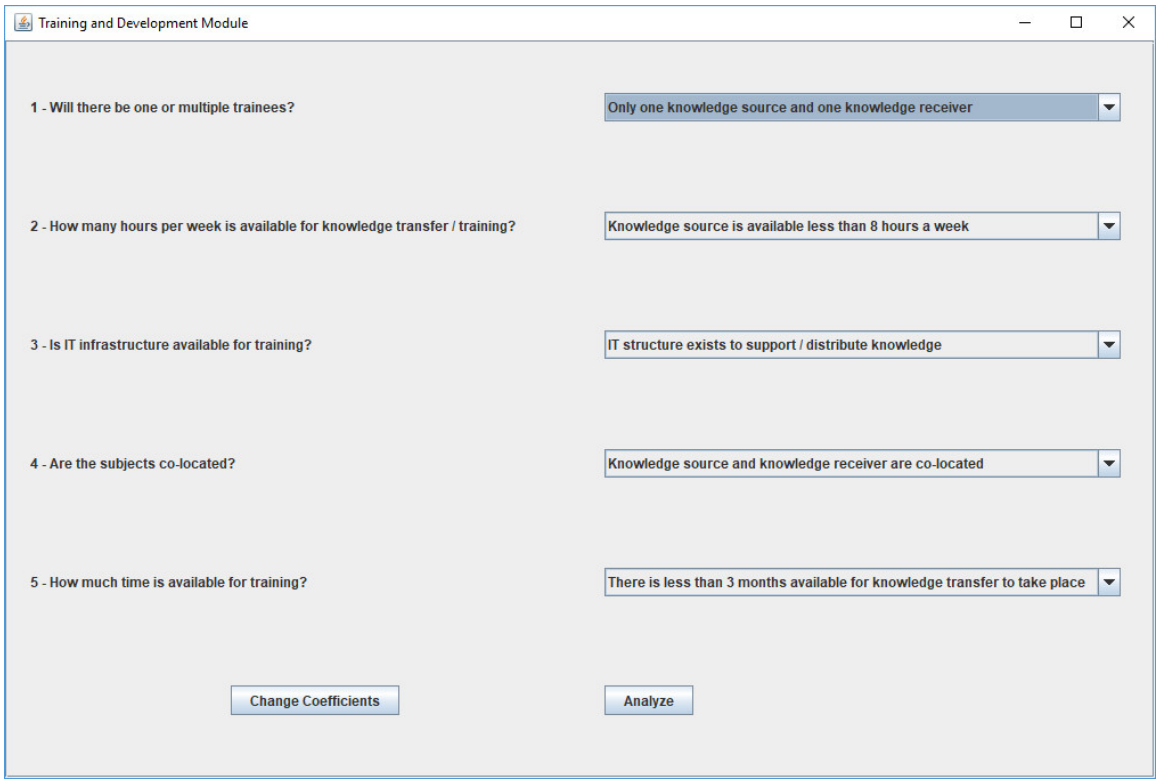


Figure A-24 – Training and Development Questionnaire

After answering the five questions regarding trainers and trainees, clicking the Analyze button allows the users to see the ratings for the 13 training and development strategies, with bar chart visualization for more intuitive results.

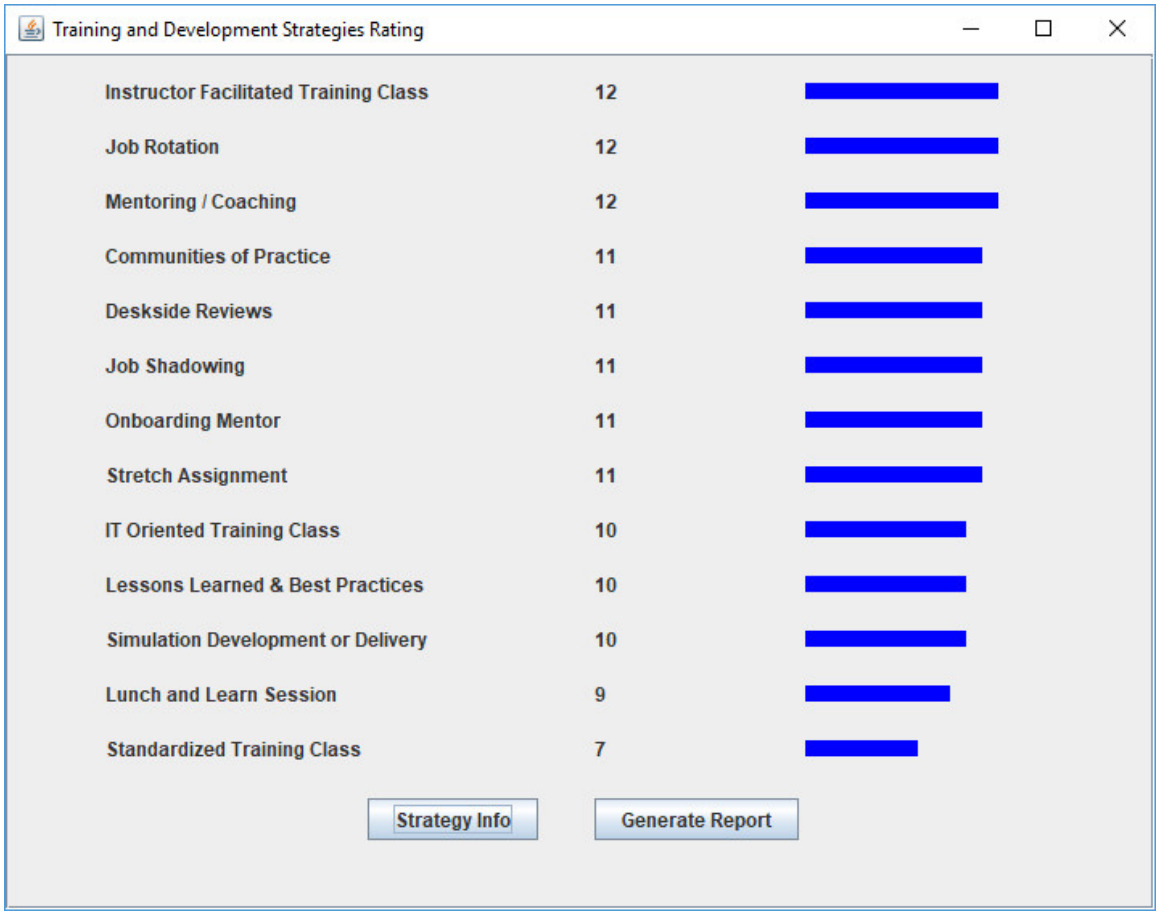


Figure A-25 – Training and Development Strategies Ranking

Strategy Info button provides users with a detailed explanation on each training and development strategy (see Figure A-26), and Generate Report button lets users to print the results for future reference.

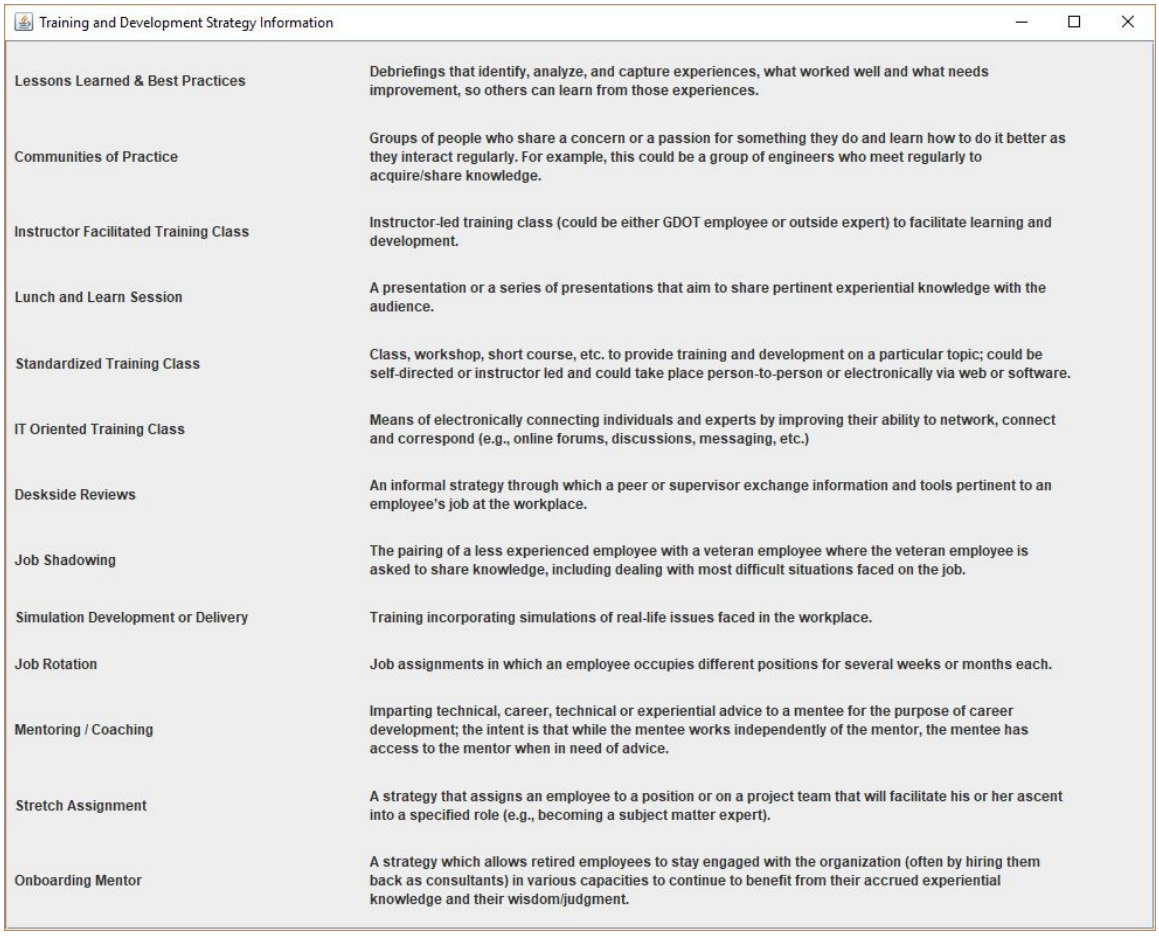


Figure A-26 – Training and Development Strategy Information

Additionally, the Change Coefficients button in the main questionnaire page enables users to change the predefined coefficients for each training and development strategy, if the user has specific knowledge or experience that would warrant the changing of the coefficients. After revising the coefficients, click the Set Coefficients button to accept the new values (see Figure A-27). The user can revert to the default values at any time by clicking the Restore Defaults button.

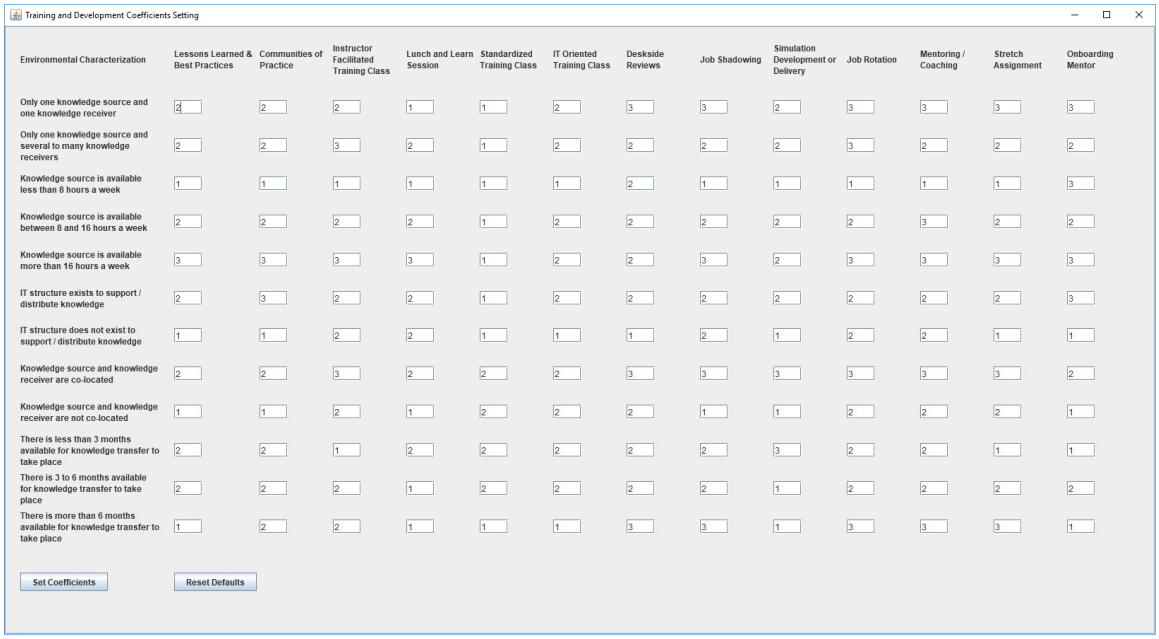


Figure A-27 – Training & Development Coefficients Setting

3. NETWORK ANALYSIS MODULE

In contrast to traditional organization charts which typically are limited to showing "who reports to whom", network analysis allows for assessment of the importance of individuals (nodes) in an organization (network) with respect to the flow of information and knowledge. The network analysis functionality of HRDT can be launched by clicking on the Network Analysis button on the landing page. The default view upon launching is the traditional organizational chart for the Office of Planning (see Figure A-28).

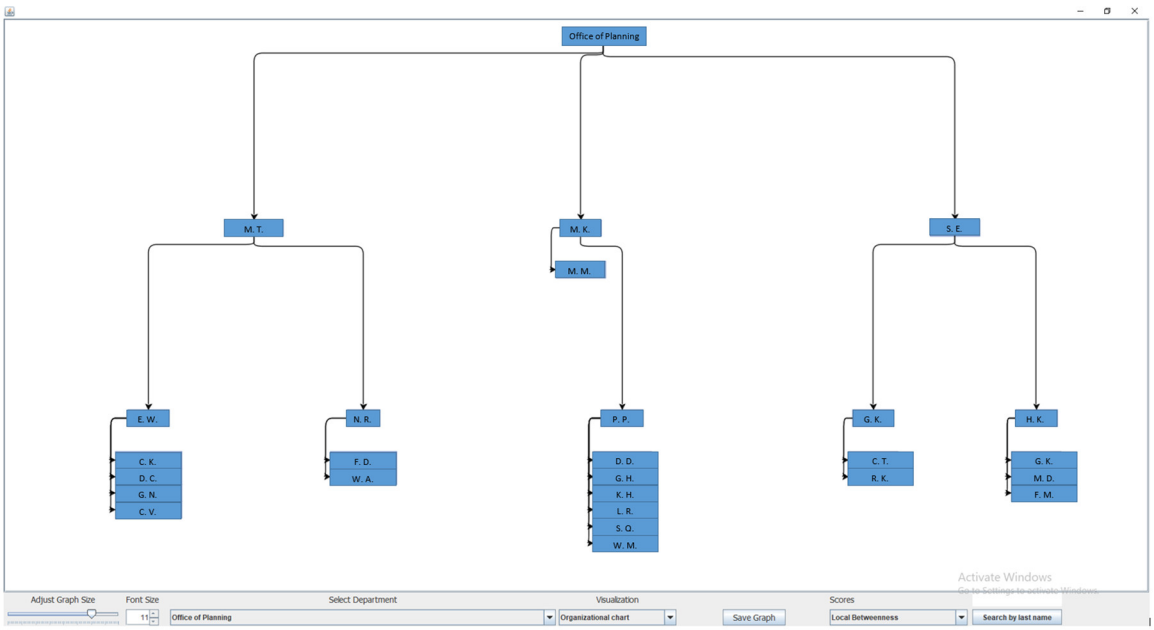


Figure A-28 – Default View of the Network Analysis Module

The bottom panel enables users to adjust quite a few features in the graph. The graph size can be adjusted by the slider, and the font size can be adjusted by the spinner (numeric up-down). To the right of the spinner is the departmental unit selector, where users can select which departmental unit's information to visualize (see Figure A-29).

Figure A-29 – Departmental Unit Selection

The users can also select how they would like to visualize the information in the selected departmental unit using the Visualization dropdown menu. The options are: Organizational chart, Network view, Hierarchical chart, Circle view and Tree chart. For example, Figure A-30 shows the Circle View for the Office of Planning.

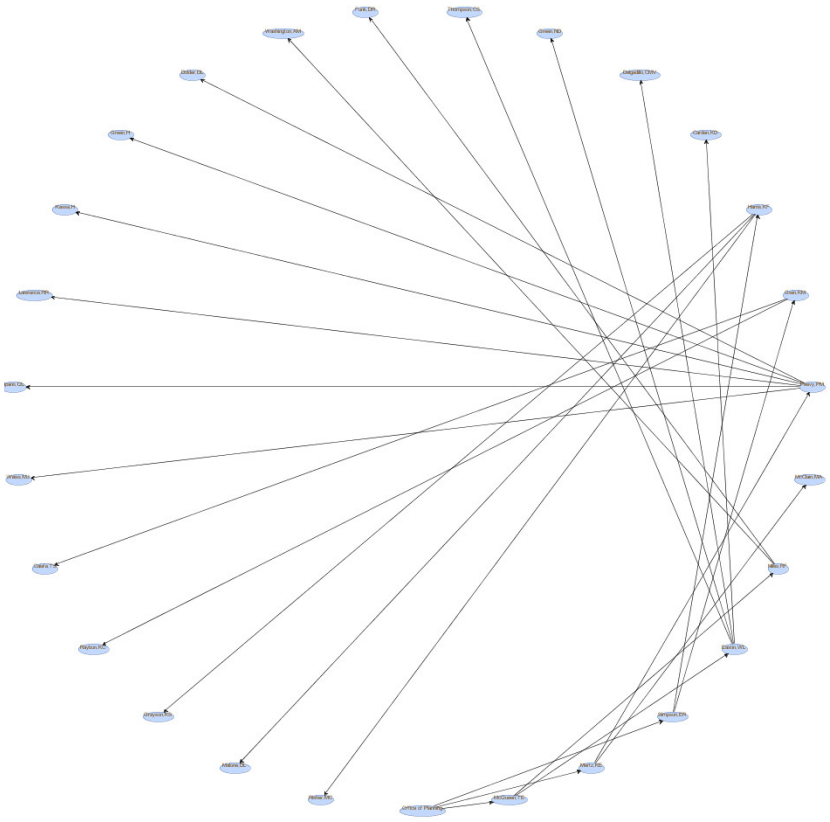


Figure A-30 – Example of Circle View

When the Network View option is selected, users can further choose the network evaluation and scoring criteria to be displayed with the Scores dropdown. The options include: Local and Global Betweenness, Local and Global Page Rank, Basic Local and Global Uniqueness, Network Local and Global Uniqueness, Cosine Dissimilarity, and Local and Global Position Level. The definition and usage of these options is provided in the main report. As an example, Figure A-31 shows the Local Betweenness centrality for the Office of Planning.

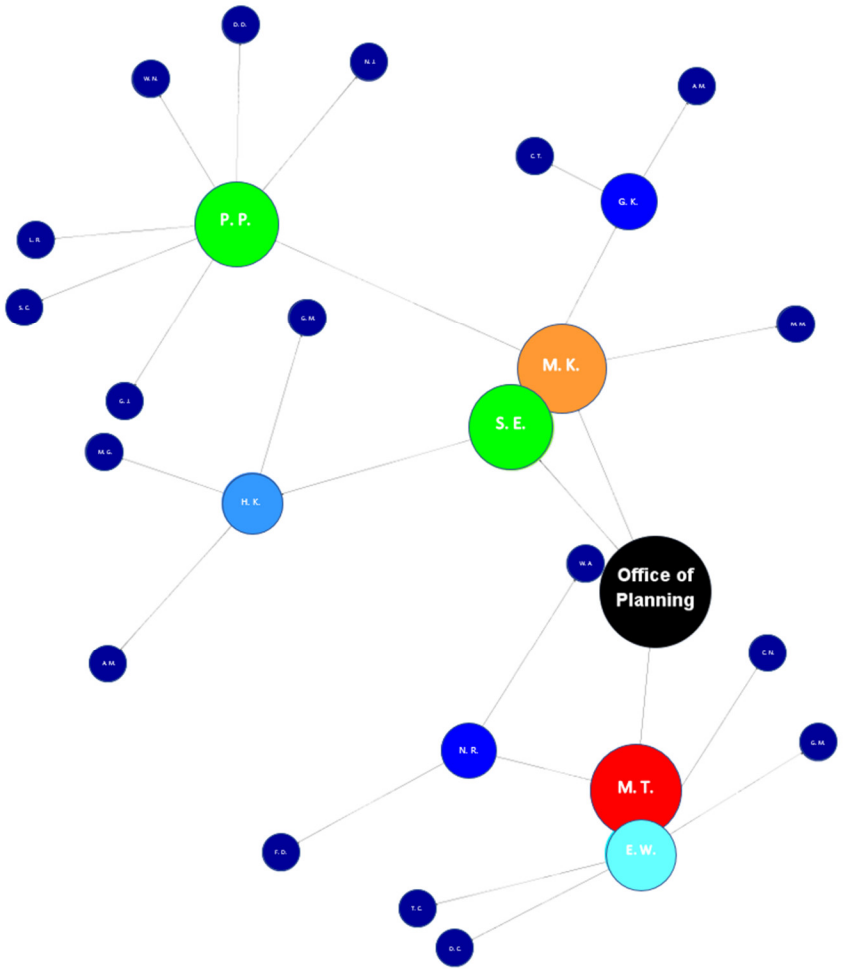


Figure A-31 – Network Visualization – Local Betweenness

In comparison, Figure A-32 shows the Local PageRank centrality for the Office of Planning. The differences between the two centralities can be seen by comparing the two outputs. In simple terms, betweenness centrality emphasizes the high-level managers, while PageRank emphasizes the mid-level managers (who are arguably more critical to the day-to-day operations of a departmental unit).

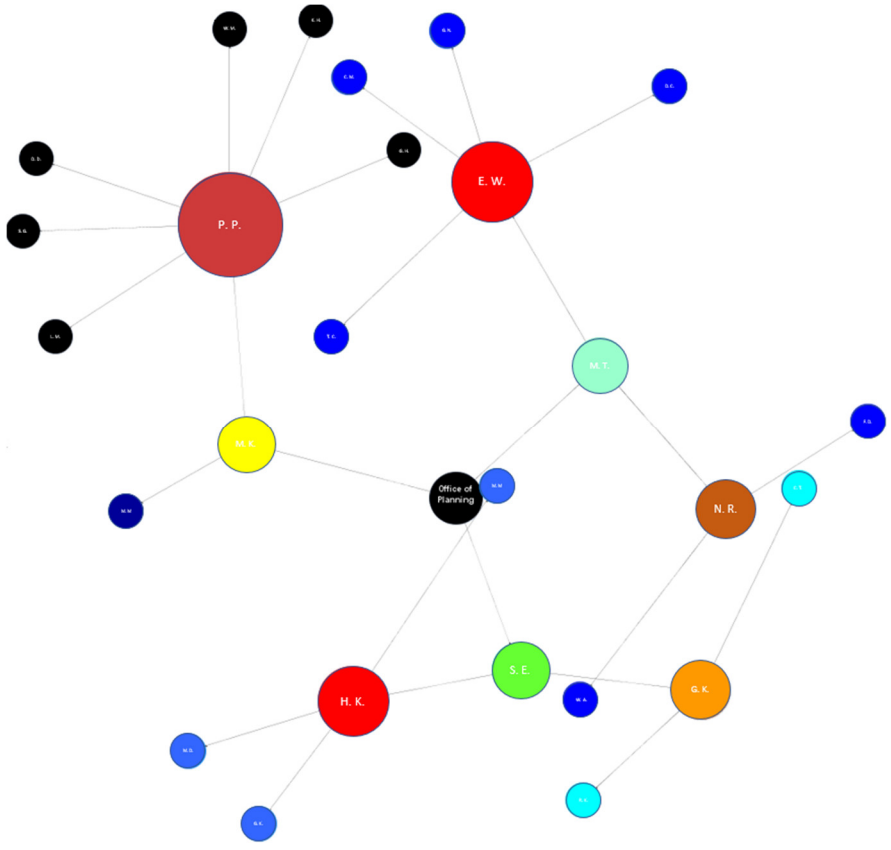


Figure A-32 – Network Visualization – Local PageRank

To the very right of the bottom panel, users may type into the text box the last name of an employee to search in the selected department. By clicking the "Search by last name" button, the graph automatically highlights the identified employees (if there are any), and zooms to the identified person (see Figure A-33).

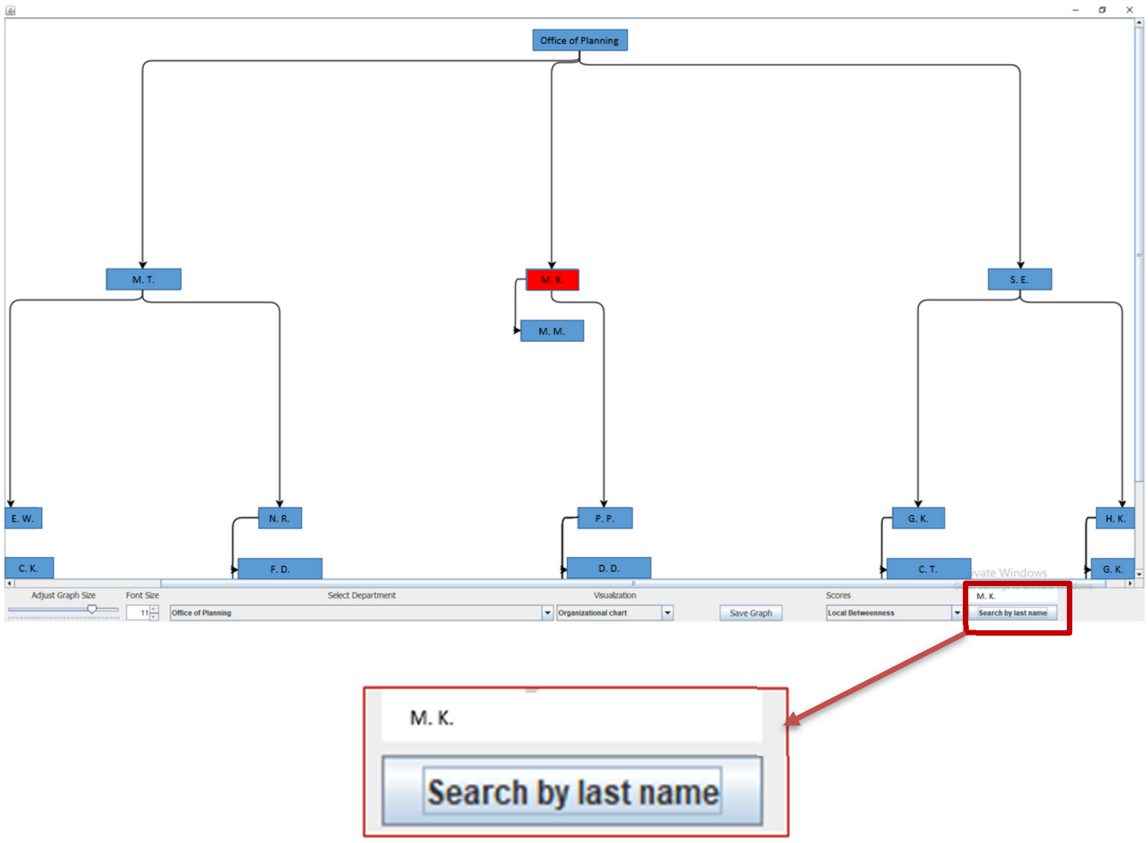


Figure A-33 – Search by Last Name in Network Analysis Module

Last but not the least, users can save any graph by clicking the Save Graph button to save the visualization to an image (.PNG) file.

4. DECISION MAKING MODULE

The Decision Making module is an integration of HRDT and Microsoft Power BI (which is provided as part of the installer package). Users may launch this module by clicking the Decision Making button on the main landing page of HRDT. The following interface (Figure A-34) will appear.

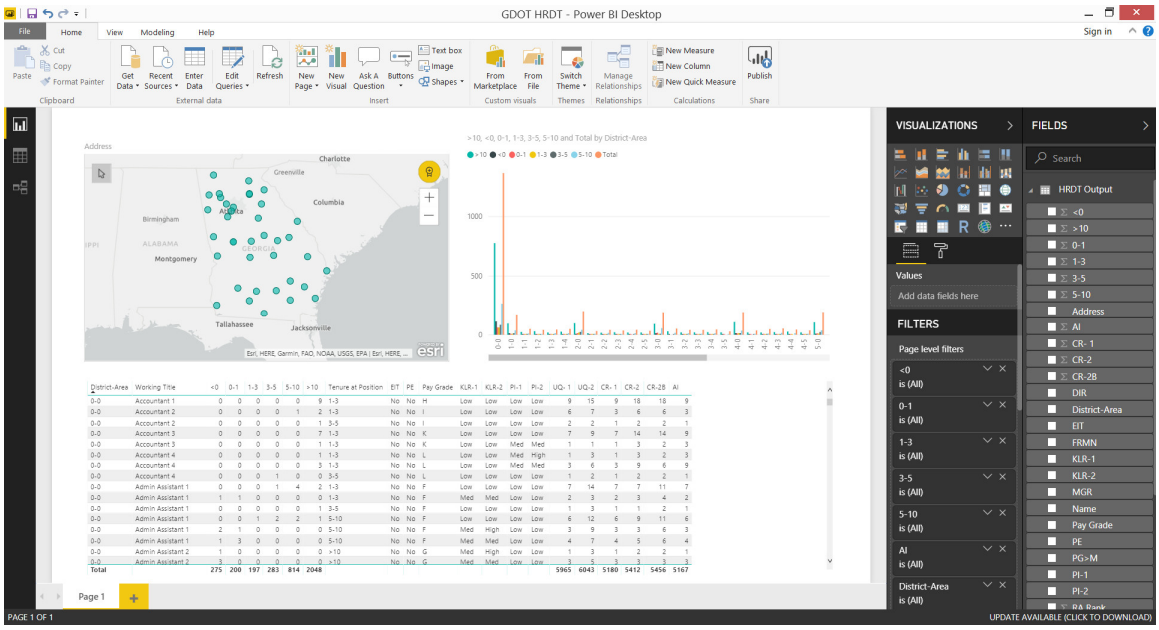


Figure A-34 – Decision Making with Microsoft Power BI

The Decision Making module consists of three main panels: geographic information system panel – referred to as GIS panel (top left), pivot chart panel (top right), and pivot table data panel (bottom). In the GIS panel, each green dot represents a district-area location. In the pivot chart, the attrition risk data are shown, with each column representing the total number of employees who are expected to retire in <0 year (i.e., should have already retired based on age and/or tenure but haven't done so yet), 0-1 year,

1-3 years, 3-5 years, 5-10 years, and >10 years. The pivot table panel shows all the attributes listed in Table 3-8 of the main report.

The auxiliary panels on the right side of the module contain the fields associated with each main panel, as well as the various filters. The default fields and filters can be modified by the user if desired; however, it is recommended that users become familiar with Power BI first before attempting to edit these.

In Figure A-34, data are shown for the entire organization. On the GIS panel, user could simply select one district-area by clicking on one of the green dots, or select multiple dots by using the rectangular select function. When one or more district-area locations have been selected, the pivot chart and the pivot table are also updated automatically to show data associated only with the selected district-areas. Similarly, if the user interacts with the pivot chart panel, the other two panels are updated accordingly to reflect the selection.

Decision Making module also has powerful filtering functions. These can be accessed from the filtering area adjacent to the main visualization page, under Page Level filters (meaning that the filters apply to all visuals on the page). Using these filters, the user can perform many powerful what-if scenario analyses. As an example, selecting the "<0" filter, setting the Filter Type to "Basic Filtering", and choosing "1" (keeping in mind "1" has been assigned to each employee who has already met the retirement requirements based on age and/or tenure, but has not yet retired) will display all those who meet the filter criteria, and the panels on the left are updated accordingly. In this example, it can be seen that 275 employees met this criterion, and 3542 did not (see Figure A-35).

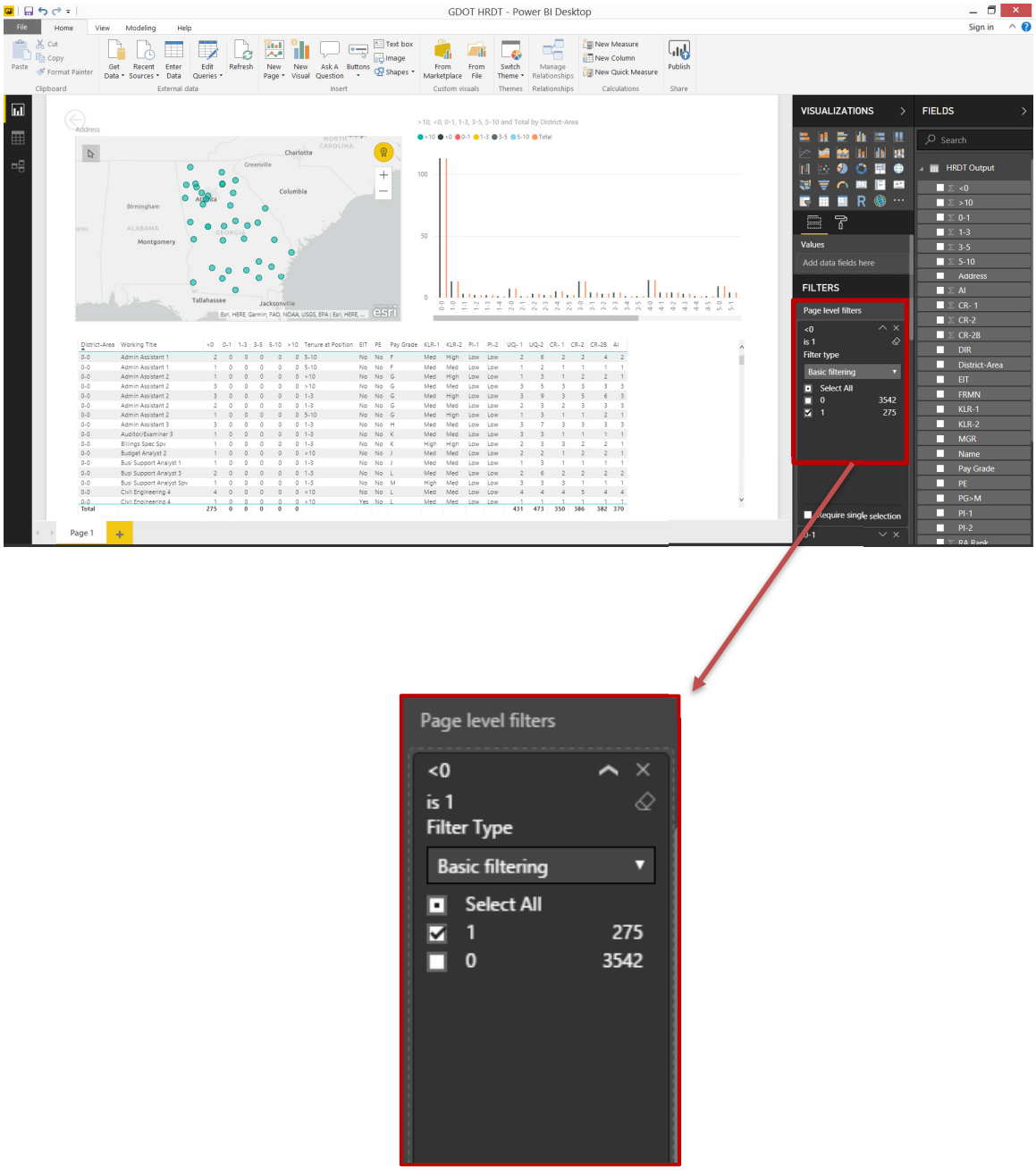


Figure A-35 – Example of Filtering Using the Decision Making Module

Continuing this example, while leaving the "<0" filter as "1", an additional filter can be imposed by selecting the "KLR-1" filter, setting the filter type as "Basic Filtering", and choosing "High" for knowledge loss risk. The data are then further filtered, to show both spatially and in tabular format, the employees who satisfy both the "<0" criteria and the "KLR-1" criteria (see Figure A-36). In this case, there are 21 such employees, and they are located in District-Area 0-0, 1-4, 2-0, 2-1, 2-2, 3-0, 3-1, 3-3, 4-3, 5-0, 5-3, 6-0 and 7-3.

Lastly, let's further assume that the user wishes to identify employee(s) who meet both of the above criteria, but also satisfy the other definition of KLR (KLR-2). This can be achieved by selecting the "KLR-2" filter, setting the filter type as "Basic Filtering", and choosing "High" once again. This means that those employees who are deemed as having a high knowledge risk using both definitions will be identified (see Figure A-37). It can be seen that only 7 such employees remain, and they are located in District-Area 0-0, 2-0, 5-0 and 6-0.

As this practical example demonstrates, by identifying such employees and their locations, HR personnel can initiate succession planning activities to ensure that the knowledge of these employees is not lost due to attrition. Similarly, running such what-if scenarios can be used as a powerful decision making tool at both the departmental unit and the organizational levels.

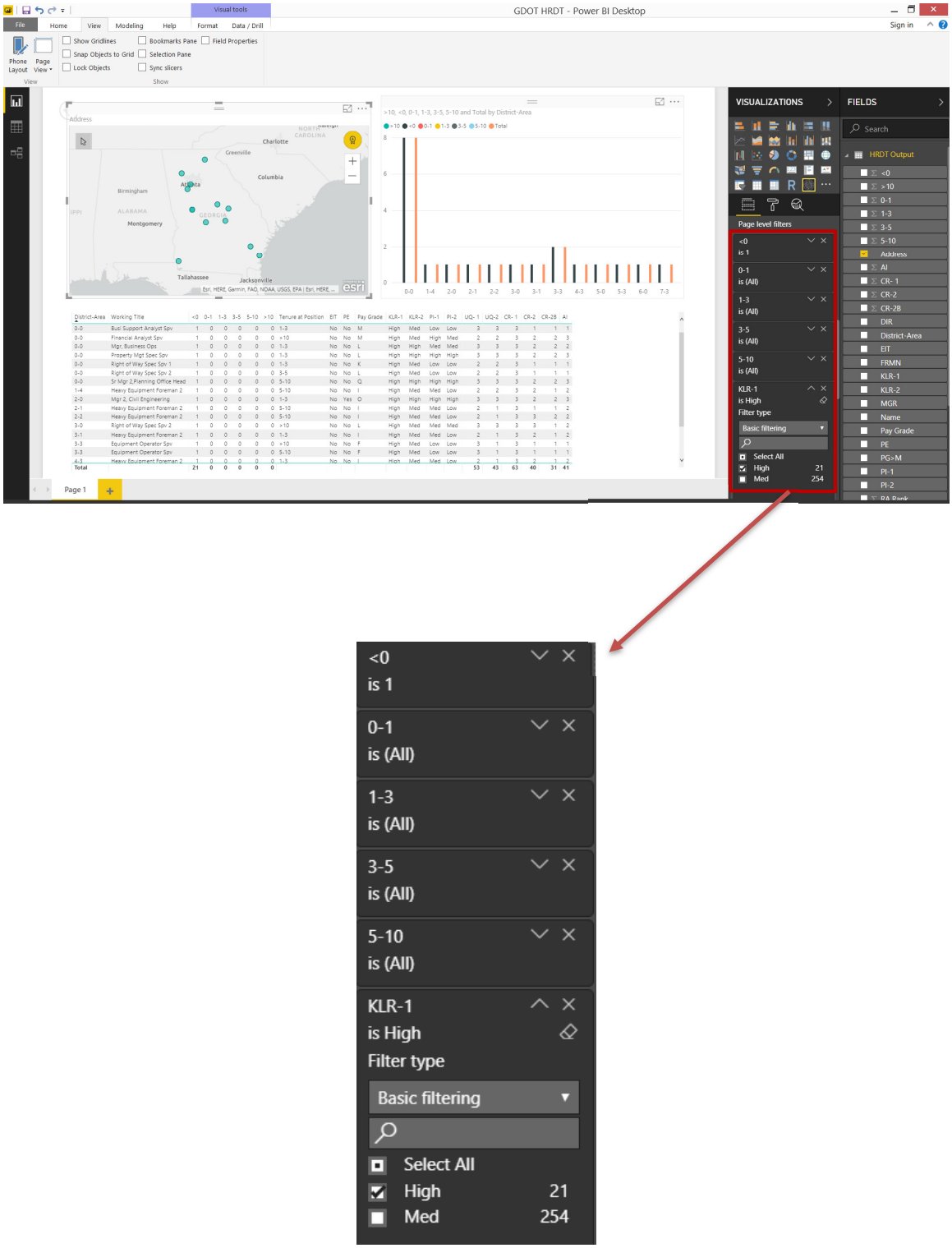


Figure A-36 – Example of Filtering Using the Decision Making Module (cont.)

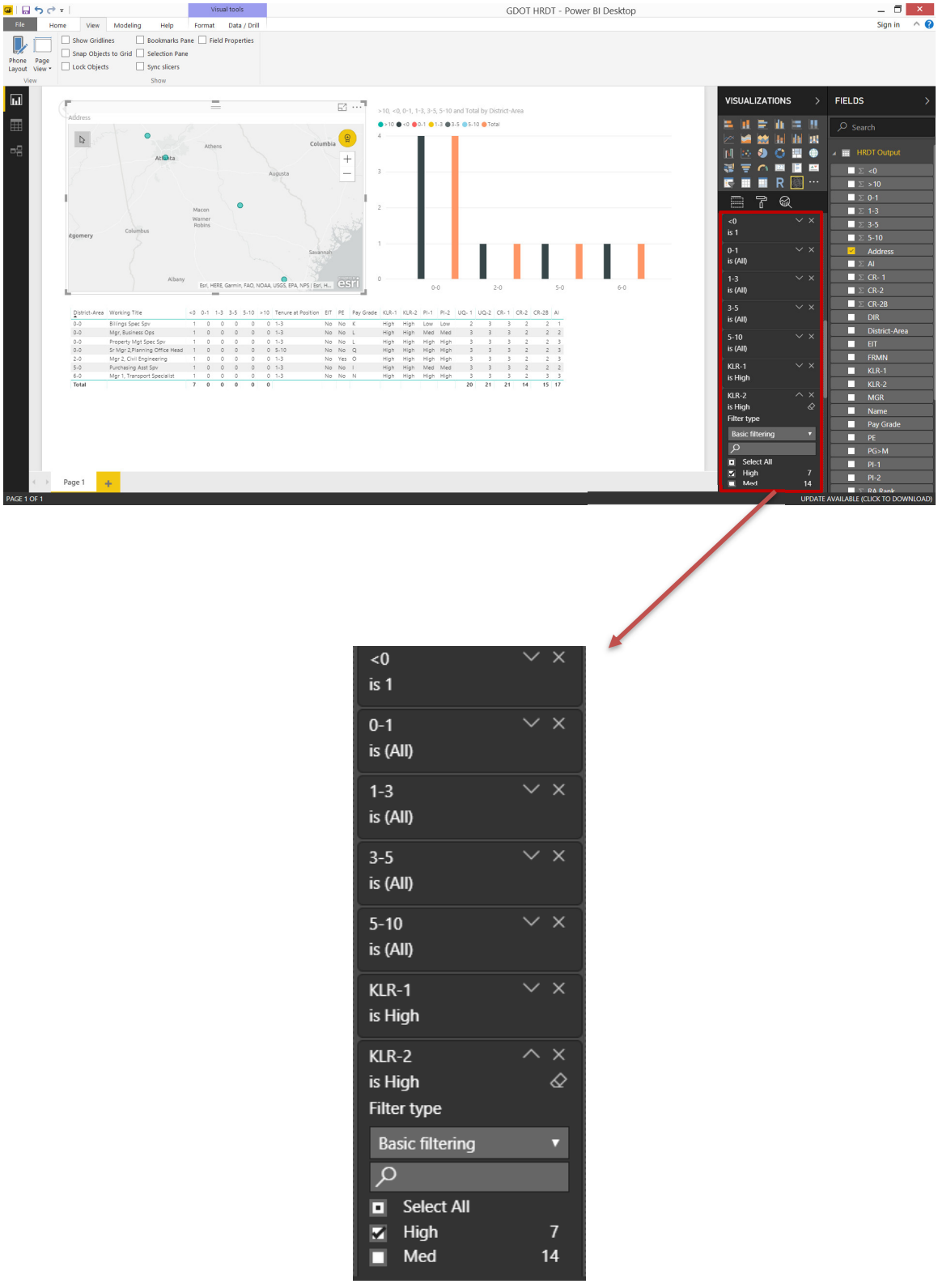


Figure A-37 – Example of Filtering Using the Decision Making Module (cont.)

After using the Decision Making module, select the "Don't save" option (see Figure A-38) in order to ensure that the program can start in its original configuration for the next user. If the user wishes to save a specific analysis performed using the module, click "Save As" from the File menu and save a separate file to the desired location on the hard drive.

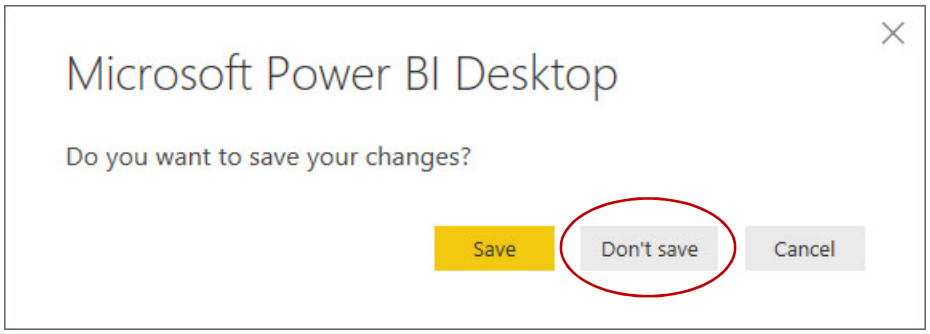


Figure A-38 – Power BI Desktop Save Dialog Box

APPENDIX B –

GDOT HR Data Tool Training Module

HR Data Tool (HRDT): A Modular System for Supporting GDOT Human Resource Planning and Decision Making

Training Module

Developed by
Georgia Institute of Technology
For
Georgia Department of Transportation

Problem Statement

- GDOT is an information rich organization; however, the existing information and data are not always rigorously analyzed. This includes the existing GDOT employee database. This problem was identified during the previously completed JOB SEEKER project.

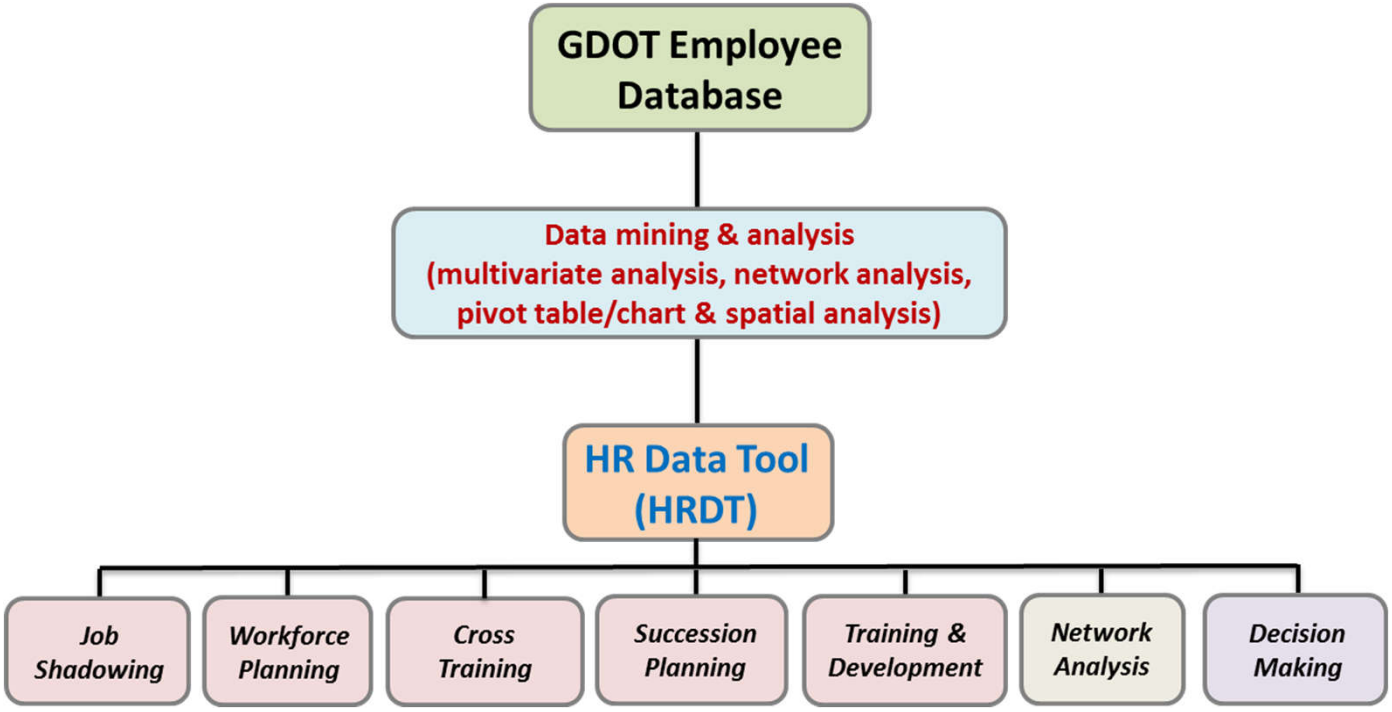
- Building upon the JOB SEEKER project, HRDT was created to help GDOT make better use of its employee data.

- HRDT is a Java-based software which provides tools for HR personnel to better analyze employee data. It includes modules for multivariate analysis, network analysis, and spatial analysis and decision-making.

- HRDT can be used for planning and strategic decision-making.

HRDT Overview

- HRDT uses the existing GDOT employee database as input, performs data mining, multivariate analysis, network analysis, and pivot table/chart and spatial analysis, culminating in 7 distinct modules:



HRDT Overview

Employee Database Structure

- Excel spreadsheet format, containing 35 columns and approximately 4,000 rows

- Data mining is performed to extract information and convert it to knowledge

- Database can be updated as frequently as needed, which in turn will update the calculations in HRDT (by clicking the “Update” button on the main page)

- **It is very important that the database structure is maintained during updates (see Main Report, Section 3.1); otherwise, HRDT will not function as intended!**

HRDT Modules

Decision Making

Network Analysis

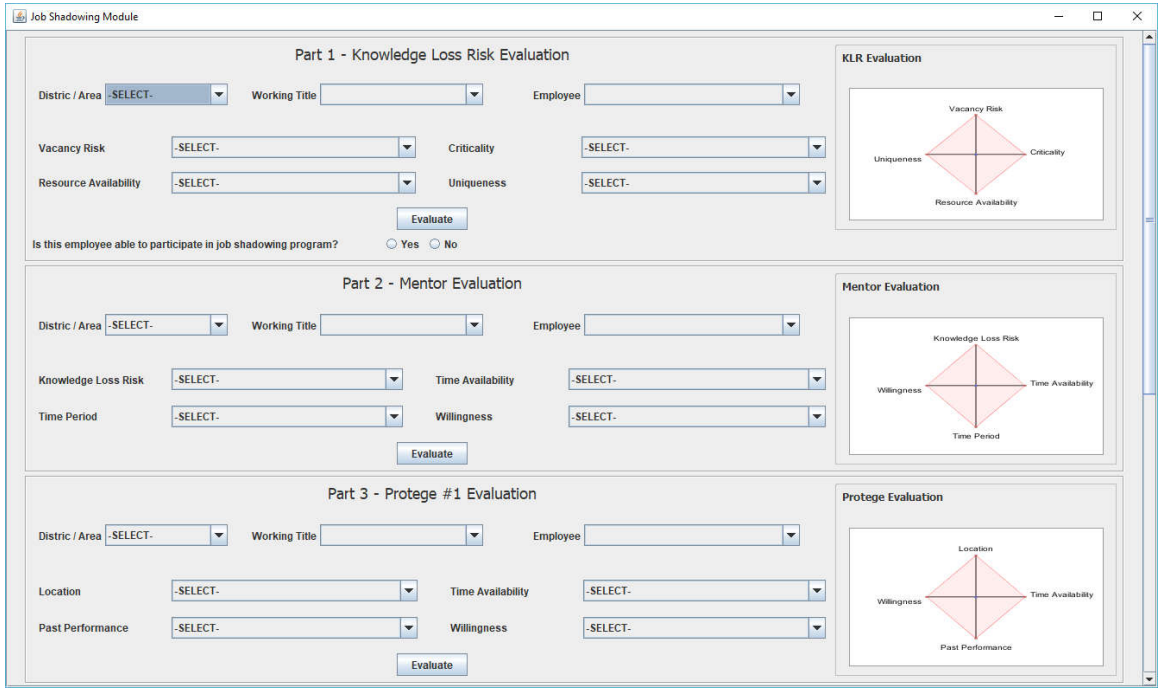
Multivariate Analysis*:

- Job Shadowing
- Workforce Planning
- Succession Planning
- Cross Training
- Training & Development

*Multivariate analyses use a consistent framework adopted from the JOB SEEKER project

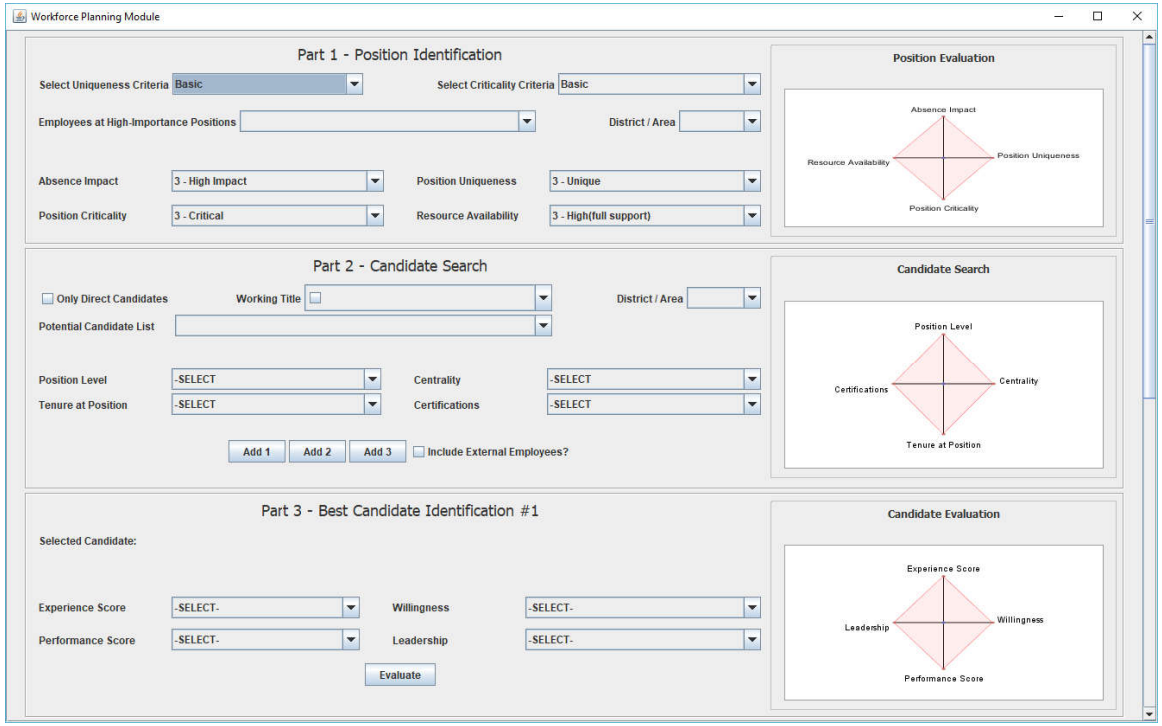
Job Shadowing

- Helps HR personnel evaluate an employee's knowledge loss risk, referred to as KLR, as well determine the most suitable mentor and protégé for participation in the job shadowing program



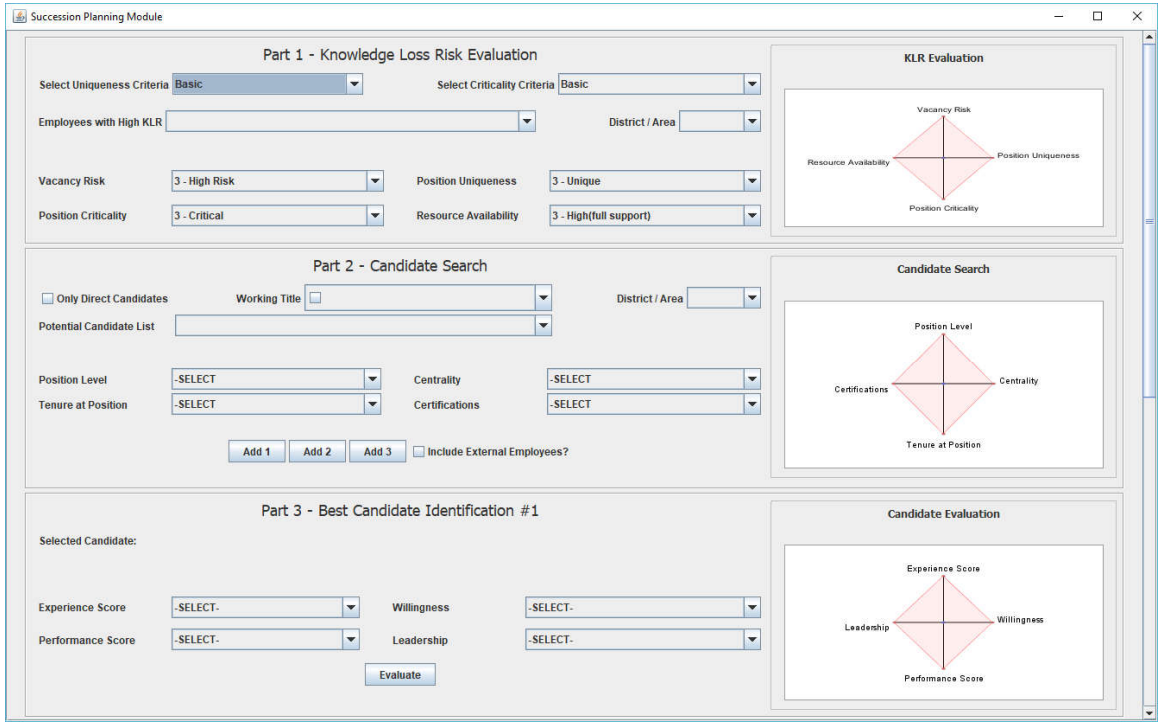
Workforce Planning

- Helps HR personnel identify high-importance positions and the employees at those positions, and helps evaluate potential candidates for workforce planning



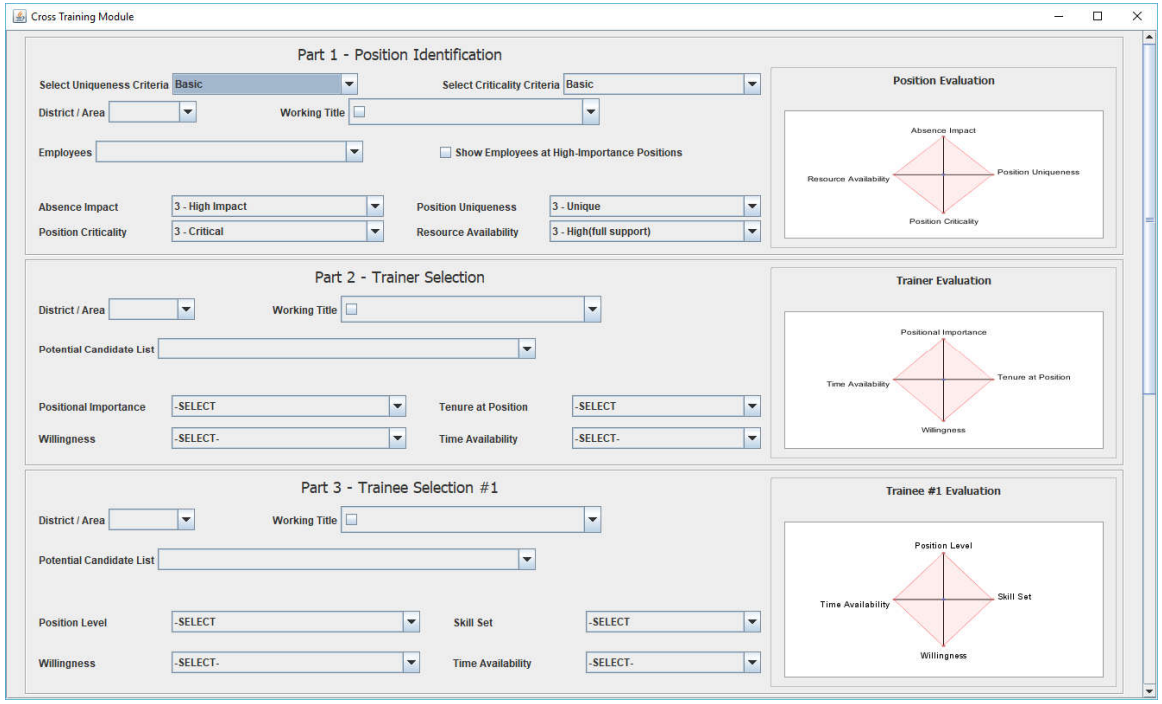
Succession Planning

- Helps HR personnel identify employees with high knowledge loss risk (KLR), and helps evaluate suitable candidates for a selected position for succession planning purposes



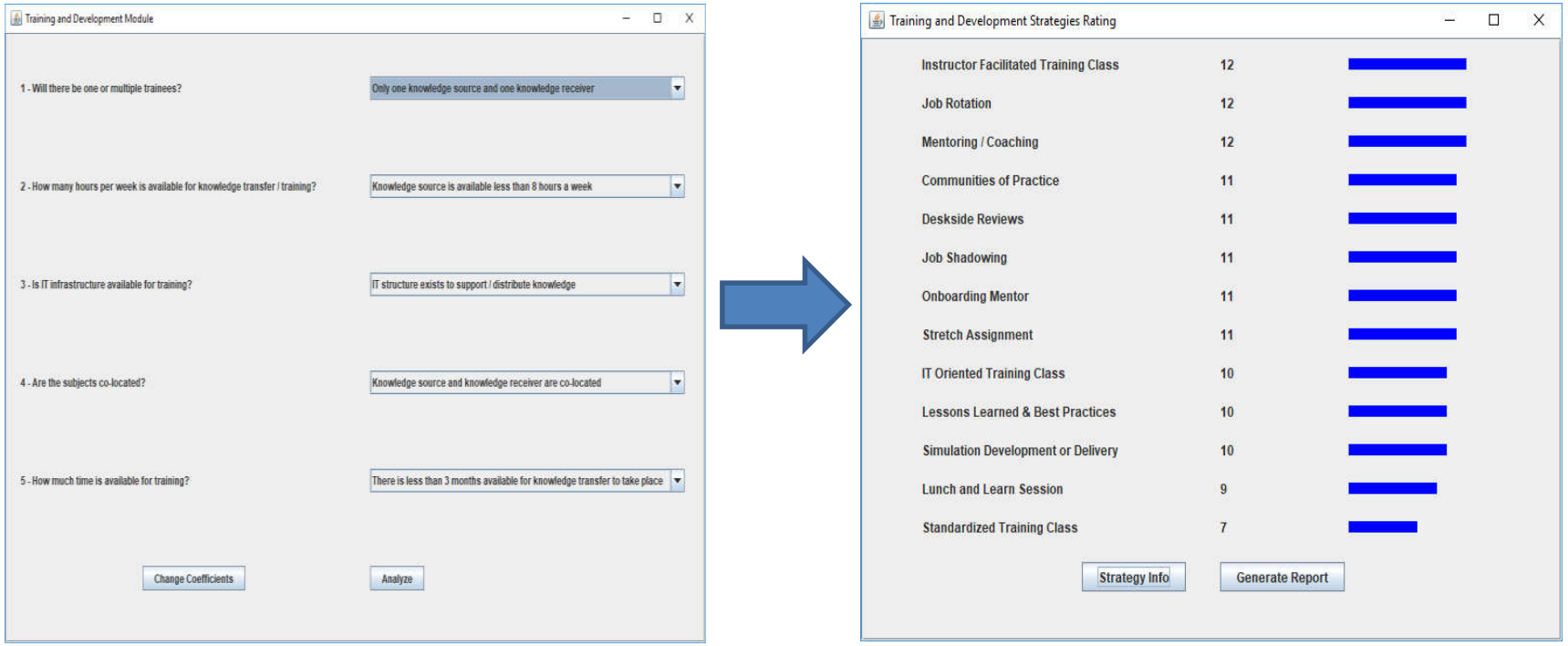
Cross Training

- Helps HR personnel identify high-importance positions and the employees at those positions, and helps evaluate suitable trainer and trainee candidate for cross-training



Training & Development

- Helps HR personnel evaluate the suitability of various training and development strategies while considering different environmental characteristics



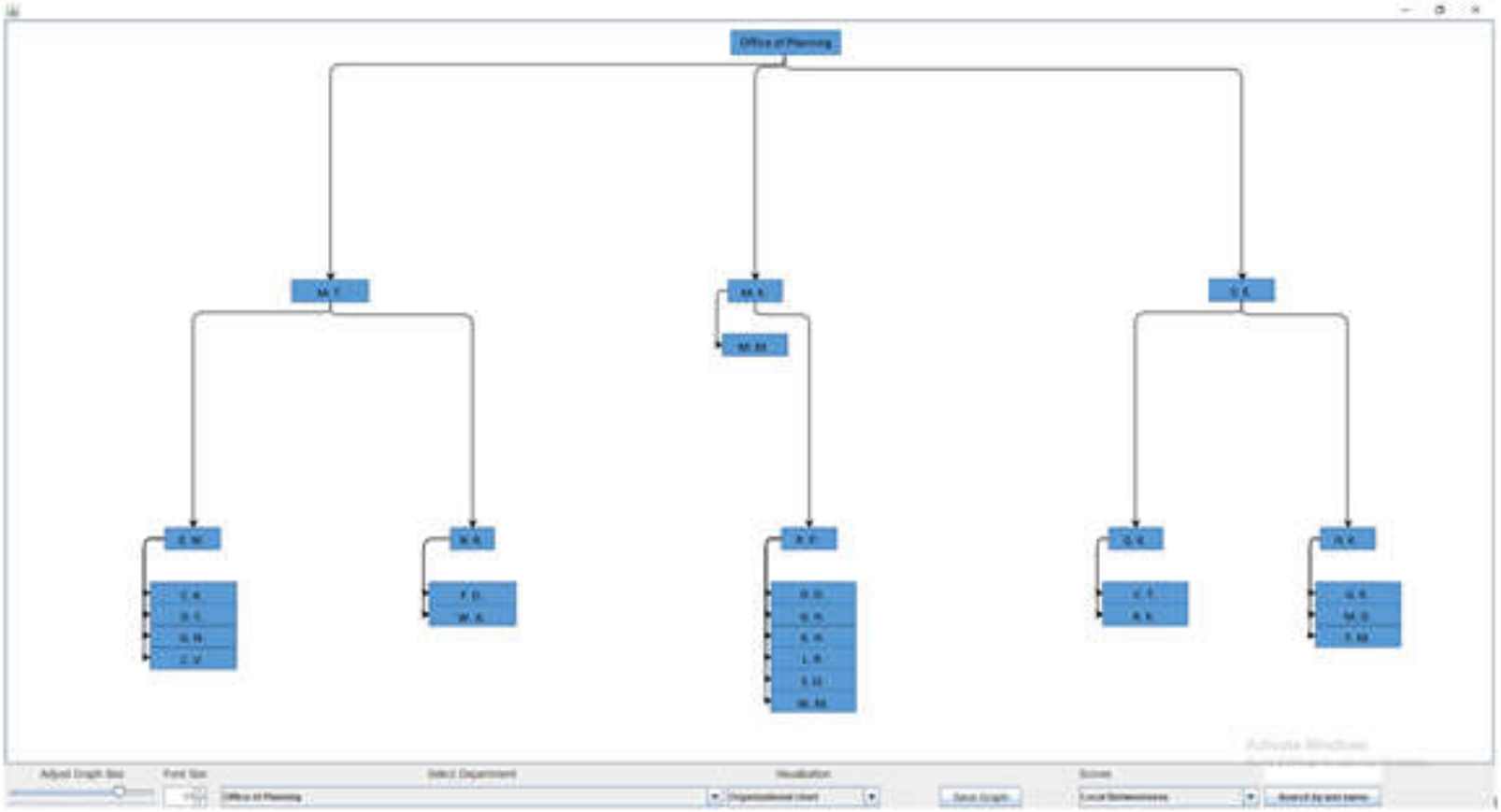
Network Analysis

- The module can reconstruct traditional organizational charts for 58 of the pre-defined departmental units in the organization (as obtained from the Master Organizational Chart), based on analysis of data extracted from the database.

- The organization chart can also be viewed in different formats (circle view, hierarchical view, tree view, etc.).

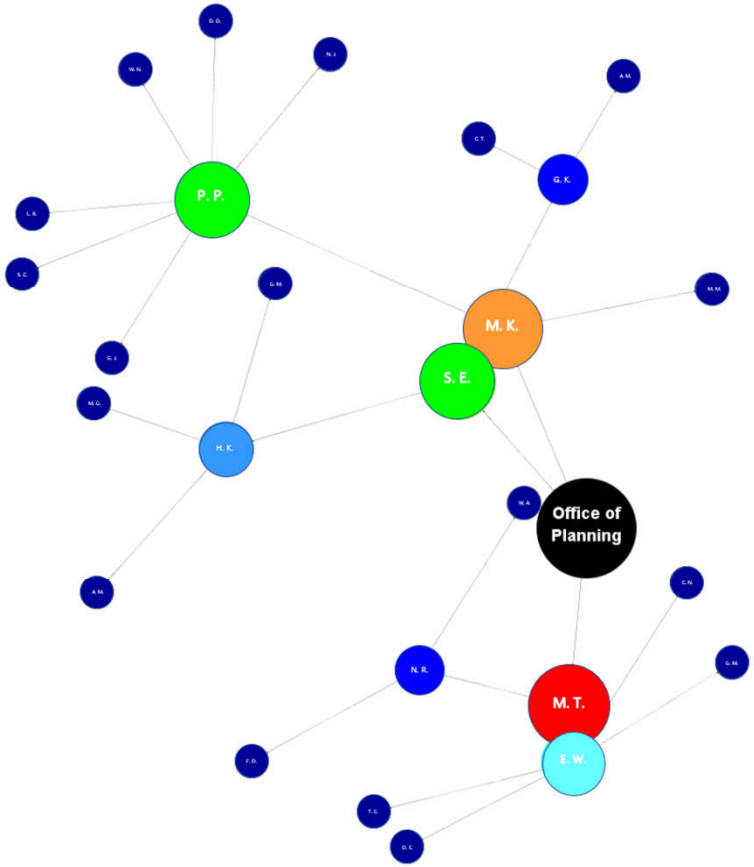
- Alternatively, the results can be viewed in the context of various network analysis techniques to assess the importance of individuals (nodes) in the organization (network) with respect to the flow of information and knowledge.

Network Analysis

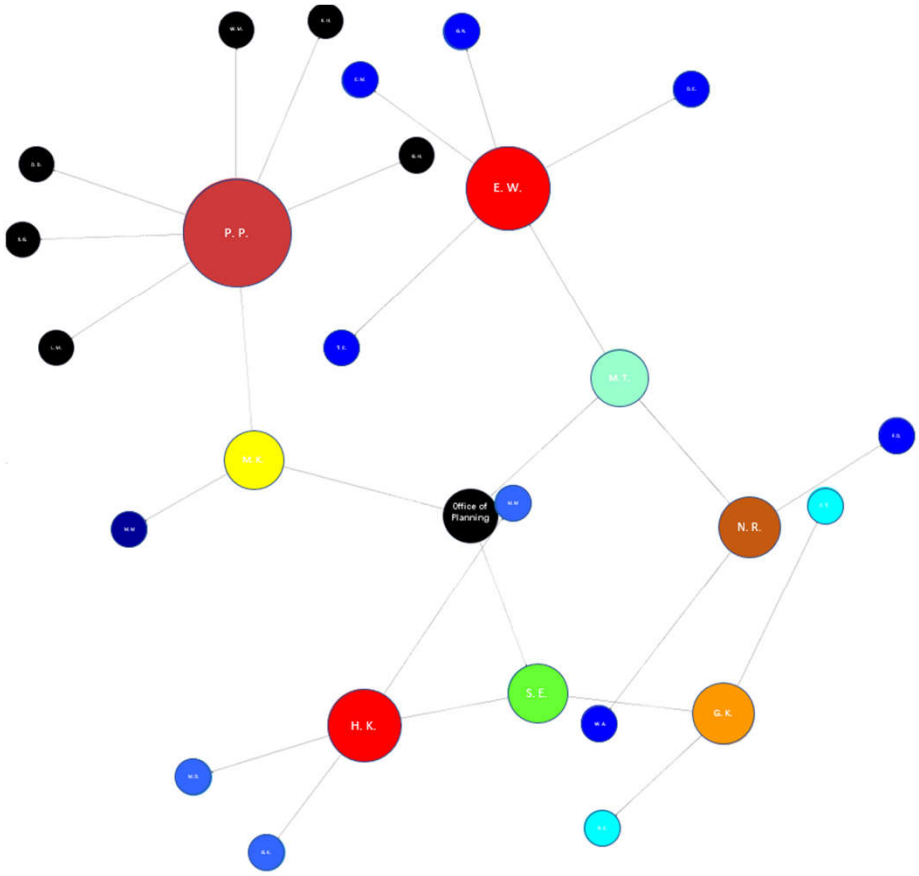


Traditional Organizational Chart View

Network Analysis



Network View
(Betweenness Centrality)



Network View
(PageRank Centrality)

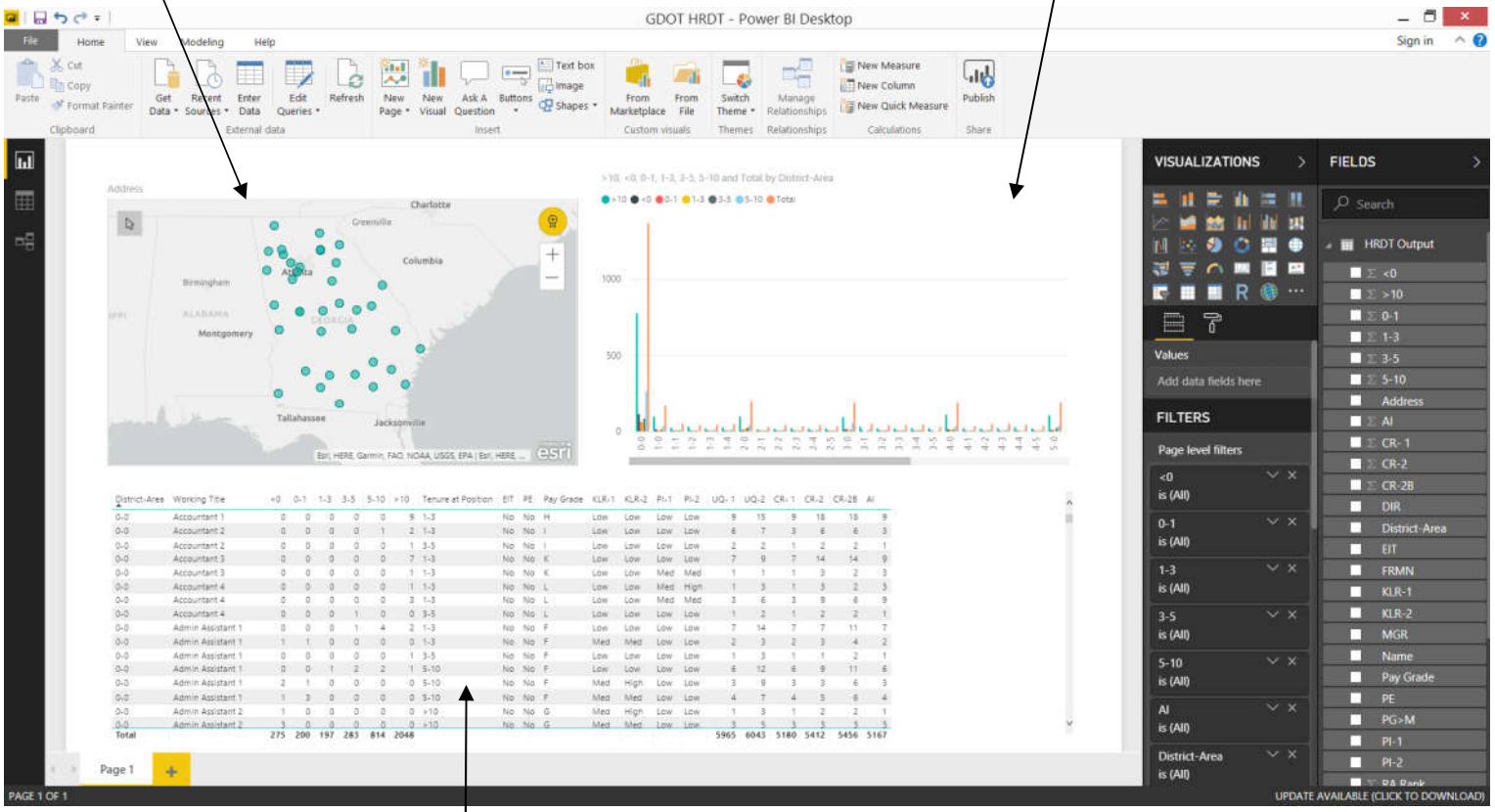
Decision Making

- The Decision Making module is an integration of HRDT and Microsoft Power BI
- Includes three main panels:
  - Geographic Information System (GIS) panel
  - Pivot chart panel
  - Pivot table panel
- Panels are inter-linked; selecting data in one panel updates all
- Provides powerful filtering and visualization capabilities to perform “what-if” scenarios

Decision Making

GIS Panel

Pivot Chart Panel



Pivot Table Data Panel

Report & User's Manual

- Main Report
  - Provides theoretical background for the framework associated with the software
  - Contains a scoring matrix used for the multivariate analyses (tasks, factors, the weights associated with each factor, etc.)
- User's Manual
  - Appendix A of main report
  - Provides instructions on how to operate each module
- Training Module
  - Appendix B of main report
  - Provides overview of HRDT to future/new users