Human Resources-Based Organizational Data Mining (HRODM): Themes, Trends, Focus, Future

Hila Chalutz-Ben Gal

1 Introduction

In recent years, there has been a trend in many organizations toward data-driven decision-making in various aspects of business (Holsapple et al. 2014) with the use of big data in daily activities (Chong and Shi 2015). Many organizations are experiencing a period of transformation as modern businesses both exploit opportunities and face numerous and complex challenges. Today’s organizational data mining (hereafter ODM) transformation is a direct result of rapid changes within organizations caused by the combined forces of demographics, globalization, and information technology. Some departments (e.g., human resources) rely on data to execute activities that were traditionally performed in a somewhat intuitive manner. This transformation plays a crucial role in firms’ ability to achieve a competitive advantage in today’s challenging economy (Kapoor and Sherif 2012; Sparrow 2012; Fulmer and Ployhart 2013). In light of the rapid changes in technology and the environment, traditional organizational metrics have become unsuitable for many situations (Fink 2010; Handa and Garima 2014; Sharif 2015).

ODM is defined as leveraging data mining (hereafter DM) tools and technologies to enhance organizational decision-making process by transforming data into valuable and actionable knowledge to gain a strategic competitive advantage (Nemati and Barko 2002, 2003). ODM domains are wide in scope. Some focus on customer relationship management, customer segmentation, retention and attrition management, risk forecasting, and profitability analysis (Kharb 2019; Meghyasi and Rad 2020). Additional ODM domains include organizational processes and Human Resources data (hereafter HRODM) for improved organizational decision-making.
This chapter focuses on HRODM utilizing data to improve people-related organizational decision-making processes. HRODM is sometimes referred to in terms such as “people analytics,” “human capital analytics,” or “human resources analytics,” among others. Within the ODM domain, HRODM is defined as “the application of sophisticated DM and business analytics techniques to the field of human resources” (Vihari and Rao, p. 1). It is also referred to as quantitative and qualitative data and information management that aims to gain insight and support decision-making processes with regard to managing people in organizations (Fitz-Enz 2000; Handa and Garima 2014; Zhao and Carlton 2015). A third definition pertains to “processes to collect, transform and manage key people related data and documents; to analyze the gathered information using DM models; and to disseminate the analysis results to decision makers for making intelligent decisions” (Kapoor and Sherif, p. 1626).

HRODM has several goals. The first is “to gather and maintain data for predicting short and long term trends in the supply and demands of workers in different industries and occupations and to help global organizations make decisions relating to optimal acquisition, development and retention of their human capital” (Kapoor and Sherif, p. 1627). The second is “to provide an organization with insights for effectively managing employees in order to achieve business goals quickly and efficiently” (Davenport et al. 2010; Hota and Gosh, p. 169). Third, some scholars emphasize that the goal of HRODM is to positively influence the successful execution of an organization’s strategy (Heuvel and Bondarouk 2016; Huselid 2015; Kapoor and Sherif 2012; Levenson 2005, 2011; Zang and Ye 2015).

In this chapter, we propose a new definition for the adoption of HRODM by focusing on the return on investment (hereafter ROI) gained by an organization when utilizing HRODM tools. ROI is a performance measure used to evaluate the efficiency of an investment or compare the efficiency of a number of different investments. ROI tries to directly measure the amount of return on a particular investment, relative to the investment’s cost. To calculate ROI, the benefit (or return) of an investment is divided by the cost of the investment. The result is expressed as a percentage or a ratio. We propose an ROI-based focus of HRODM, because it enables organizational insights and supports decision-makers with respect to the human capital dilemma by providing business insight and consequently helping them make better business decisions. Our proposed ROI-based approach is grounded upon a systematic review and analysis of the literature in the field. In recent years, the connection between ODM and HR has resulted in a growing body of literature that proposes various approaches to combining the two disciplines, sometimes in an unstructured, blunt manner. Moreover, despite notable evidence of a growing interest in HRODM, researchers have found very limited scientific evidence to help decision-makers determine whether and how to adopt and implement HRODM (Rasmussen and Ulrich 2015).

This chapter aims to bridge this gap by proposing an ROI-based review of HRODM in the sense that the efforts required to adopt analytic data mining methods and apply them to HR tasks must be justified. This chapter has two objectives. The first objective is to provide an integrative analysis of the literature on the
topic of HRODM through the lens of ROI to provide scholars, executives, and practitioners with a comprehensive but practical view of the topic (Huselid 2015). The chapter emphasizes the developments in HRODM in recent years, particularly by highlighting works that have been published within the past 5 years (Vihari and Rao 2013; Rasmussen and Ulrich 2015; Heuvel and Bondarouk 2016; Bamber et al. 2017). The second objective is to systematically analyze the literature from the ROI perspective, highlighting scientific evidence to assist decision-makers in determining how to adopt HRODM (Rasmussen and Ulrich 2015). This work aims to aid both researchers and practitioners with respect to specific directions within HRODM in which an expected ROI may be found.

Understanding what we have learned and how it has changed the ODM field helps direct future research. To this end, this chapter asks and answers three interrelated research questions (Cuozzo et al. 2017):

**RQ1.** What are the major themes that have been developed within HRODM research?

**RQ2.** What are the focus and ROI-based critique of HRODM research?

**RQ3.** What is the future of HRODM research?

This chapter includes four sections. The methodology section outlines the database development approach. The results and discussion sections answer the first two research questions through descriptive statistics and a critique of the results from categorizing the HRODM literature. This section also discusses how we developed and applied the ROI theoretical framework. In the third section, we answer the last research question by discussing key implications for scholars and practitioners and noting a few directions for future research. Finally, in the fourth section, we utilize two real-life examples to demonstrate HRODM implementation.

## 2 Methodology

The methodological approach for the review and analysis comprised four steps. First, we developed a database by undertaking a comprehensive and systematic search to identify and extract all the relevant literature on HRODM that has been published in peer-reviewed academic journals. Second, in an iterative process between theoretically derived and empirically emerging themes, we developed a template for analyzing the reviewed articles (Table 1). Third, a manual content analysis of the retrieved articles, based on the template, was used to extract descriptive and qualitative conceptual data. Finally, the results were interpreted and the findings meaningfully synthesized (Short 2009; Webster and Watson 2002). This method was used to ensure a comprehensive, meaningful, and high-quality data compilation (Cuozzo et al. 2017).
Table 1 Classification system for analyzing HRODM articles

<table>
<thead>
<tr>
<th>Code</th>
<th>Cluster/category</th>
<th>Number of articles</th>
<th>%</th>
<th>Example references</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quantitative</td>
<td>4</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>E2</td>
<td>Qualitative</td>
<td>6</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>E3</td>
<td>Mixed methods</td>
<td>4</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Conceptual</td>
<td>36</td>
<td>45</td>
<td>Davenport et al. (2010), Kapoor (2010), Wiblen et al. (2010), Harriss et al. (2011), Snell (2011), Minbaeva and Collings (2013), and Pape (2016)</td>
</tr>
<tr>
<td>C1</td>
<td>Management tools</td>
<td>10</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>General</td>
<td>18</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>Specific</td>
<td>8</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>CB1</td>
<td>General</td>
<td>4</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>CB2</td>
<td>Specific</td>
<td>7</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>Informative</td>
<td>9</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>Specific</td>
<td>5</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>Literature review</td>
<td>3</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>HRA trends</td>
<td>2</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

Based on: Chalutz Ben-Gal (2019)

2.1 Database Development

The initial step comprised the identification of the relevant research. To capture previously published research, we used 11 EBSCO online databases.1 We conducted a Boolean search using “human resources analytics” as a key search term within the

1Databases included for the review: Business Source Premier; EconLit; Regional Business News; SocINDEX; ERIC; Library, Information Science & Technology Abstracts; Historical Abstracts; Communication & Mass Media Complete; GreenFILE; Political Science Complete; PsycARTICLES.
In reviewing and analyzing the selected papers, we identified four HRODM research clusters: empirical, conceptual, case-based, and technical. These research clusters are depicted in Fig. 1. This categorization is useful in developing an ROI-based analysis of HRODM (Webster and Watson 2002; Gilbert et al. 2008; Bukhari et al. 2017).

### Classification System

The articles were first coded by the lead category (i.e., cluster) and then checked for consistency by an external judge who had extensive experience with the topic. Any discrepancies were reviewed and discussed before a final classification was agreed upon. Rather than describe each category in the framework as presented here in Table 1, we outline each at the beginning of the corresponding discussion in the descriptive results (Cuozzo et al. 2017).

### Results and Discussion

In this section, we use descriptive statistics and commentary to answer the first two research questions: RQ1. What are the major themes that have been developed within HRODM? RQ2. What are the focus and ROI-based critique of HRODM? The

---

2 Additional search terms included “organizational data mining,” “workforce analytics,” “people analytics,” and “human capital analytics.”
Fig. 2 HRODM publications over time. (Based on: Chalutz Ben-Gal (2019))

data reported in Fig. 2 and in Tables 1, 2, 3, and 4 form the basis for this section. Additionally, the discussion is complemented by further analysis that delves deeper than the descriptive results.

We analyze the findings of our systematic review of a sample of 80 articles associated with research in HRODM according to the chronological development of this research (presented in Table 2 and Fig. 2). We thereby identify shifting trends over time and extract key themes of existing HRODM literature. Additionally, we analyze and present key trends in HRODM research in Table 3 in a unique synthesis (presented below).

3.1 Emergence of HRODM Research

The results of our research, displayed in Table 2, clearly show an increasing interest in the topic of HRODM over time (see also Fig. 2). We identified three periods of HRODM research. The first is a period of incubation (2000–2005) during which 4% of the HRODM research was published. The second was a period of incremental growth (2006–2010) during which 10% of the HRODM research was published. Finally, there was a period of substantial growth (2011–2016) when 86% of the HRODM research was published. In line with this typology, and consistent with previous research (Rasmussen and Ulrich 2015; Nemati and Barko 2002, 2003), our study results demonstrate that the research attention devoted to HRODM has increased in recent years. The shift in publication over this 17-year period underscores the growing academic interest in the field of HRODM (Bose 2015; Kazakovs et al. 2015). Moreover, the understanding of HRODM has changed over time. While early publications examined HRODM from a narrow economic
Table 2  HRODM research characteristics by period of publication

<table>
<thead>
<tr>
<th>Year of publication</th>
<th>2000–2005 (6 years) Incubation period (1 ≥ publications) N = 3 (4%)</th>
<th>2006–2010 (5 years) Incremental growth (1 &lt; publications &lt; 4) N = 8 (10%)</th>
<th>2011–2016 (6 years) Substantial growth (publications ≥ 10) N = 69 (86%)</th>
<th>2000–2016 Total (17 years) N = 80 (100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of journal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRM</td>
<td>2(3)</td>
<td>4(5)</td>
<td>32(40)</td>
<td>38(48)</td>
</tr>
<tr>
<td>Management &amp; Business</td>
<td>0</td>
<td>4(5)</td>
<td>30(37)</td>
<td>34(43)</td>
</tr>
<tr>
<td>Engineering</td>
<td>0</td>
<td>0</td>
<td>2(3)</td>
<td>2(3)</td>
</tr>
<tr>
<td>Other</td>
<td>1(1)</td>
<td>0</td>
<td>5(6)</td>
<td>6(7)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Research cluster</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empirical</td>
<td>2(3)</td>
<td>1(1)</td>
<td>11(14)</td>
<td>14(18)</td>
</tr>
<tr>
<td>Conceptual</td>
<td>1(1)</td>
<td>3(4)</td>
<td>32(40)</td>
<td>36(45)</td>
</tr>
<tr>
<td>Case Based</td>
<td>0</td>
<td>2(3)</td>
<td>9(11)</td>
<td>11(14)</td>
</tr>
<tr>
<td>Technical</td>
<td>0</td>
<td>2(3)</td>
<td>17(21)</td>
<td>19(24)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Geographical region</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td>1(1)</td>
<td>3(4)</td>
<td>49(61)</td>
<td>53(66)</td>
</tr>
<tr>
<td>Europe</td>
<td>2(3)</td>
<td>2(3)</td>
<td>8(10)</td>
<td>12(15)</td>
</tr>
<tr>
<td>Asia</td>
<td>0</td>
<td>2(3)</td>
<td>8(10)</td>
<td>10(13)</td>
</tr>
<tr>
<td>Africa/Middle East</td>
<td>0</td>
<td>1(1)</td>
<td>4(5)</td>
<td>5(6)</td>
</tr>
</tbody>
</table>

Based on: Chalutz Ben-Gal (2019)
Values = Number of articles; values in brackets = % of articles

<sup>a</sup>Adds to 101% due to percentage rounding
Table 3  Trends in HRODM research

<table>
<thead>
<tr>
<th>Trend</th>
<th>Challenges &amp; outcomes</th>
<th>ROI</th>
<th>Example references</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evidence-based approach in HRODM</strong></td>
<td>Adoption of correct tool</td>
<td>High</td>
<td>Bassi (2011), Nemati and Barko (2002, 2003), McIver et al. (2018), and Strohmeier (2018)</td>
</tr>
<tr>
<td><strong>HRODM as decision-making support tool</strong></td>
<td>Various analytical techniques</td>
<td>High</td>
<td>Pessach et al. (2020), Dulebohn and Johnson (2013), Singh and Roushan (2013), Holsapple et al. (2014), Rasmussen and Ulrich (2015), Pape (2016), and Chamorro-Premuzic et al. (2017)</td>
</tr>
<tr>
<td><strong>HRODM as management fad</strong></td>
<td>HRODM is not part of DM</td>
<td>Low</td>
<td>Rasmussen and Ulrich (2015) and Nemati and Barko (2003)</td>
</tr>
</tbody>
</table>

Based on: Chalutz Ben-Gal (2019)

perspective by highlighting technical aspects (e.g., Lazear 2000), the relevance of HRODM has gained importance both in research and in practice from a strategic and managerial perspective, which has transformed it into a vibrant and interesting topic of research.

More specifically, HRODM research has evolved such that in the incubation period (2000–2005), none of the publications found their way into management nor business journals, whereas almost 40% (37%) of the publications did so in the substantial growth period (2011–2016). Moreover, the study results indicate that a vast share of HRODM research (91%) was published in either HR management or in management and business journals. Forty-eight percent of HRODM research was published in HR management journals, while 43% of HRODM research was published in management and business journals.

These findings indicate the increase in the strategic importance of the field. One explanation is the growing centrality of human capital as a key organizational asset (Bontis and Fitz-Enz 2002; Fitz-Enz 2000; Nemati and Barko 2002, 2003). Both HR and ODM as a broader field are in a constant state of change (Bamber et al. 2017; McIver et al. 2018). A second explanation is the growing availability of readily accessible data, which can be transformed into valuable and actionable insights through the implementation of ODM tools (Macan et al. 2012; Strohmeier 2018). These findings also show that our ROI-based analysis is an appropriate platform to expand upon in order to determine precisely how management HRODM.

Research results suggest an emerging shift over time regarding the geographical regions upon which HRODM research focuses. Most articles on the topic of HRODM that specified a geographical region in the substantial growth period shifted from Europe (10% of publications) to North America (61% of publications).
Table 4  ROI-based analysis of HRODM

<table>
<thead>
<tr>
<th>Study Authors</th>
<th>Research Cluster</th>
<th>Logic</th>
<th>Analytics</th>
<th>Measurements</th>
<th>Processes</th>
<th>ROI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harrison and Getz</td>
<td>Empirical</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>High</td>
</tr>
<tr>
<td>Hou</td>
<td>Empirical</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>High</td>
</tr>
<tr>
<td>Ramamurthy et al.</td>
<td>Empirical</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>High</td>
</tr>
<tr>
<td>Sharif</td>
<td>Empirical</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>High</td>
</tr>
<tr>
<td>Bose</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td></td>
<td>×</td>
<td>Medium</td>
</tr>
<tr>
<td>Church et al.</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>Medium</td>
</tr>
<tr>
<td>Huselid</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>High</td>
</tr>
<tr>
<td>Levenson</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>High</td>
</tr>
<tr>
<td>Momin and Mishra</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td></td>
<td>×</td>
<td>Medium</td>
</tr>
<tr>
<td>Newcomer and Brass</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>Medium</td>
</tr>
<tr>
<td>Perrin</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td></td>
<td>×</td>
<td>Medium</td>
</tr>
<tr>
<td>Rasmussen and Ulrich</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>Medium</td>
</tr>
<tr>
<td>Sharma et al.</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>Medium</td>
</tr>
<tr>
<td>Steffi et al.</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td></td>
<td>×</td>
<td>Low</td>
</tr>
<tr>
<td>Ulrich and Dulebohn</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td></td>
<td>×</td>
<td>Medium</td>
</tr>
<tr>
<td>Zang and Ye</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>Medium</td>
</tr>
<tr>
<td>Zhao and Carlton</td>
<td>Conceptual</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>High</td>
</tr>
<tr>
<td>Frigo et al.</td>
<td>Case Based</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Russell and Bennett</td>
<td>Case Based</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td>Medium</td>
</tr>
<tr>
<td>Chong and Shi</td>
<td>Technical</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Karasek</td>
<td>Technical</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Kazakovs et al.</td>
<td>Technical</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>Low</td>
</tr>
<tr>
<td>Korpela</td>
<td>Technical</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Stone et al.</td>
<td>Technical</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td>Medium</td>
</tr>
<tr>
<td>Welbourne</td>
<td>Technical</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td>Low</td>
</tr>
</tbody>
</table>

Based on: Chalutz Ben-Gal (2019)

N = 25. Included in 2015 publications analysis only (represents the highest publishing year in the Substantial Growth period)

*aBoudreau and Ramstad (2006)*

This focus on North America could be linked to the emerging trend, which originated in the United States, of linking technology and data along with the major effect that technology has on organizations as a whole (Chamorro-Premuzic et al. 2017).

Most of the research articles are conceptual (45%) rather than purely technical (24%). The conceptual studies in HRODM provide management and analytical tools to facilitate working processes and procedures. They include talent analytics (Burdon and Harpur 2014), tools for improved organizational decision-making (Minbaeva and Collings 2013; Pape 2016), and a conceptual framework (Boudreau and Ramstad 2006). The focus has thereby shifted over time from a predominance of conceptual articles to technical articles, which comprise nearly one-quarter (24%) of the total number of articles (see Table 2). To a certain extent, this may be due to
the growing interest in specific topics within ODM (Macan et al. 2012; Yadav 2014; Momin and Mishra 2015).

3.2 Trends in HRODM Research

Our integrative review reveals that HRODM research is dominated by four trends (see Table 3). We synthesize these trends by depicting their key challenges, outcomes, and ROI.

The first identified trend in HRODM research is the exploration of HRODM as a strategic management tool. This approach yields a high ROI for the organization because its impact may be on the organization as a whole and on the business level for the purpose of continuous improvement (Delbridge and Barton 2002). Where HRODM is presumed to be an integral part of management processes, the key challenges associated with this trend include answering questions regarding specific strategic measures. One example is organizational key performance indicators (KPIs), for example: turnover and churn, which have a long-term business impact on the organization as a whole (Levenson 2005, 2015; Newcomer and Brass 2015; Welbourne 2015). Along these lines, one researched theme associated with this trend is the management and organizational interfaces within organizations (Huselid 2015; Xiu et al. 2017; McIver et al. 2018).

The second identified trend in HRODM research is the evidence-based approach to organizational data mining research. This approach also yields a high ROI for the organization because it uses a variety of methodological and technological tools to predict improved individual or organizational performance. The key challenges associated with this trend include answering key questions regarding which tool would be the correct one to adopt for a specific people analytic challenge and which form of technology to use (Strohmeier 2018).

The third identified trend in HRODM research that uses ODM for effective organizational processes involves incorporating data mining as an effective decision-making support tool (Dulebohn and Johnson 2013; Singh and Roushan 2013; Holsapple et al. 2014; Chamorro-Premuzic et al. 2017). The ROI associated with this trend is high because it suggests efficiency in the decision-making processes. The key challenges associated with this trend include the efficiency of the process itself, e.g., collecting and analyzing the data, thereby raising issues of efficiency and effectiveness (Rasmussen and Ulrich 2015; Pape 2016; Dastyar et al. 2017).

Finally, the studies focusing on the future of ODM incorporate a fourth trend in HRODM research. This approach yields a low ROI because it is speculative in nature. The key challenges associated with this trend include discussions of whether HRODM should be part of the HR function and the role of HR professionals (Rasmussen and Ulrich 2015). This paper builds upon this trend and pinpoints specific practical directions regarding how to implement HRODM. We thus move to
our substantive contribution, an ROI-based analysis of HRODM that sets the ground for our proposed future research avenues in ODM.

### 3.3 Theoretical Framework: ROI-Based Analysis of HRODM

The theoretical framework of ROI guided our analysis. The literature suggests that ROI is an important measurement tool that may assist stakeholders in managerial decision-making. ROI is rooted in early theoretical research in the accounting and management professions that aimed to provide a qualitative approach to decision-making. ROI is also used in various academic fields (Philips 2012; Bontis and Fitz-Enz 2002; Bukhari et al. 2017). One example is in the corporate training and education literature, where ROI is used to measure the impact of training and educational investments on an organization’s “bottom line,” i.e., organizational performance measures (Charlton and Osterweil 2005).

We examine the results of this study from an ROI-based perspective for two reasons. First, we believe that this framework is suitable in light of the limited high-quality research that has been conducted in the field (Fink 2010; Handa and Garima 2014; Xiu et al. 2017). Second, we believe that analyzing HRODM from an ROI-based perspective can increase the chances of the practical adoption of HRODM. We therefore categorized the research reported in this article based on the LAMP framework (Boudreau and Ramstad 2006). We identified this framework as a suitable framework to analyze ROI in the field of HRODM. In particular, the LAMP framework assists in analyzing useful components of HRODM, i.e., “logic,” “analysis,” “measurement,” and “process” (Boudreau and Ramstad, p. 27). Using this categorization, we found that the majority of articles from the empirical and conceptual research clusters resulted in high or medium levels of ROI. Additionally, we found that most studies focusing on cases or technical aspects of HRODM resulted in medium or low levels of ROI. We summarize the coding of our sample in Table 4.

### 3.4 Empirical Studies

Empirical studies attempt to obtain knowledge in the field of HRODM. The majority of studies were conducted by direct and indirect observations and/or experience (Aral et al. 2012). Their analyses were either quantitative or qualitative. An advantage of empirical studies is that by quantifying evidence or making sense of it in a qualitative manner, scholars answer empirical questions that are clearly defined and answerable based on data and the use of the evidence collected. The research designs vary by field and by the question being investigated. Some scholars perform mixed-methods research, combining qualitative and quantitative forms of analysis to better answer their research questions, especially in the social sciences.

The contributions of empirical studies in the literature are evident as they explore new and current trends in HRODM research in all or some of the following ways. First, they conduct interviews with practitioners in a variety of organizations from different industries on the topic of HRODM. Additionally, they conduct interviews with thought leaders in the area of human capital analytics and research. Finally, they attempt to draw informative conclusions in the area of HRODM (Lazear 2000; Fink 2010; Hausknecht 2014; Kandogan et al. 2014; Sharif 2015).

As documented in Table 4, a review of the literature suggests that most empirical studies apply the LAMP model (Boudreau and Ramstad 2006) in a meaningful manner, thereby yielding a high level of ROI. Moreover, our results indicate that most empirical studies are consistent with the LAMP model because they focus on at least three of the model’s components by providing meaningful content to the “logic,” “analytics,” “measurement,” or “process” of HRODM (Boudreau and Ramstad 2006; Gilbert et al. 2008). Moreover, empirical studies yield the highest ROI because they focus on organizational practices and business performance. Furthermore, empirical studies highlight the metrics used by organizations as well as the impact of HRODM on business outcomes (Lazear 2000; Lawler III et al. 2004). Finally, empirical studies tend to use strategic tools, such as forecasting techniques to predict various human-related measures (Hausknecht 2014; Bondarouk and Ruël 2013; Del Angizan et al. 2014).

To conclude, there appear to be clear benefits to exploring HRODM from an empirical standpoint. Some of the benefits include increased organizational performance, greater accuracy regarding performance specifications, accurate and rapid assessment processes, and better HR processes (Harrison and Getz 2015; Hou 2015; Ramamurthy et al. 2015).

### 3.5 Conceptual Studies

Some of the studies covered in this systematic review offer conceptual contributions to the field of HRODM. The advantage of the conceptual studies is that their contributions are wide; they range from providing management tools (Davenport et al. 2010; Wiblen et al. 2010; Kapoor 2010; Snell 2011; Harris et al. 2011) to providing an ethical perspective to talent analytics (Burdon and Harpur 2014) and adopting a data-mining-based approach (Ramamurthy et al. 2015). Their contributions relate to various content areas in the HRODM field (Gilbert et al. 2008). Some studies apply statistics, technology, and expertise to large sets of people data, which results in improved organizational decisions (Minbaeva and Collings 2013; Pape 2016). Other studies emphasize analytical processes to enhance an organization’s competitive advantage (Burdon and Harpur 2014).

Reviewing the conceptual literature, we identified four main themes: Organizational data mining’s ROI, the conceptual framework contribution, the Orga-

Interestingly, the results presented in Table 4 indicate that like the empirical studies, the majority of conceptual studies apply the LAMP model (Boudreau and Ramstad 2006) in a meaningful manner, thereby yielding a medium to high level of ROI. Moreover, our results indicate that most conceptual studies are consistent with the LAMP model because they focus on at least two of the model’s components, i.e., “logic,” “analytics,” “measurement,” or “process” (Boudreau and Ramstad 2006).

Conceptual studies in HRODM yield a medium to high ROI because some propose new frameworks to analyze and implement organizational and employee data (Davenport et al. 2010; Wiblen et al. 2010; Garcea et al. 2011), while others discuss the roles and responsibilities of HR in this transformational era of technological change and globalization (Kapoor 2011; Snell 2011; Harris et al. 2011; Burdon and Harpur 2014). Some of the reviewed literature focuses on performance management (Schläfke et al. 2012; Ding and Zhang 2014; Church et al. 2015; Ryan and Herleman 2016) and may provide a new method for managers to obtain insight into the effectiveness of employee performance and, ultimately, organizational performance (Ding and Zhang, p. 5). Some of the conceptual studies take a broader approach to the measurement of human capital in light of constant organizational change (Baron 2011; Carlson and Kavanagh 2011; Ingham 2011; Dulebohn and Johnson 2013).

The increased level of ROI that is derived from the conceptual literature on HRODM (medium to high level of ROI, as indicated in Table 4) is also derived from some of its strategic implications. Some conceptual studies in HRODM provide tools for workforce analytics and emphasize the strategic importance of HRODM within the organizational context (Huselid and Becker 2011; Van Barneveld et al. 2012; Hota and Ghosh 2013; Boudreau 2014; Guszcza and Richardson 2014; Handa and Garima 2014; Holsapple et al. 2014; Bose 2015; Huselid 2015; Rasmussen and Ulrich 2015; Ryan and Herleman 2016; Sharma et al. 2015; Steffi et al. 2015; Zang and Ye 2015; Ulrich 2016).

To conclude, the common feature of the conceptual studies is that they articulate a clear connection between the investment in analytics and organizational effectiveness. Moreover, they all have indicators of increased level of ROI. Finally, the conceptual research studies present a robust approach for strategic alignment.
with state-of-the-art organizational processes (Boudreau and Ramstad 2006), which complements their overall effectiveness.

### 3.6 Case-Based Studies

The case-based literature has two foci. First, it covers studies that provide practical examples of organizations that have implemented HRODM and recommendations for successful implementation. Second, some studies were written by scholars or practitioners who have consulting experience in the area of HRODM and share it with their readers. An advantage of the case-based studies is their practicality in the field of HRODM (Gilbert et al. 2008).

The results presented in Table 4 indicate that in contrast to the empirical and conceptual studies, most case-based studies do not apply the LAMP model (Boudreau and Ramstad 2006) in a meaningful manner, and they therefore yield a medium to low level of ROI. Moreover, our results indicate that most case-based studies are inconsistent with the LAMP model and therefore yield lower levels of ROI because they focus on only one or two of the model’s components, i.e., “logic,” “analytics,” “measurement,” or “process” (Davenport 2006; Fitz-Enz 2000; Briggs 2011; Mondore et al. 2011; Boyd and Gessner 2013; Singh and Roushan 2013; Varshney et al. 2014; Frigo et al. 2015; Russell and Bennett 2015).

To conclude, the common grounds for what we categorized as case-based studies (Gilbert et al. 2008) is that the majority do not articulate a clear connection between HRODM investment, organizational effectiveness, and ROI. Moreover, they provide limited scientific evidence to aid decision-makers concerning whether to adopt or implement organizational data mining tools within an organization (Strohmeier 2018).

### 3.7 Technical Studies

The technical literature analyzed in this study has four focus areas: The studies present informative research on the topic of HRODM (Welbourne 2015), focus on a specific subject within HRODM (Perrin 2015), present a literature review (Vihari and Rao 2013; Chong and Shi 2015), or illustrate future trends in HRODM (Yadav 2014; Momin and Mishra 2015). Thus, the advantage of technical studies is their specificity (Gilbert et al. 2008).

The results presented in Table 4 indicate that similar to the case-based literature, and in contrast to the empirical and conceptual studies, the majority of technical studies do not apply the LAMP model (Boudreau and Ramstad 2006) in a meaningful manner, and they therefore yield a medium to low level of ROI. Moreover, our results indicate that most technical studies are inconsistent with the LAMP model because they focus on only one or two of the model’s components, i.e., “logic,”
“analytics,” “measurement,” or “process” (Mayo 2006; Rivera and Smolders 2013; Stone and Dulebohn 2013; Vihari and Rao 2013; Fernández-Delgado et al. 2014; Yadav 2014; Chong and Shi 2015; Karasek 2015; Kazakovs et al. 2015; Korpela 2015; Momin and Mishra 2015; Perrin 2015; Stone et al. 2015; Ulrich and Dulebohn 2015; Welbourne 2015; Ryan and Herleman 2016).

To conclude, the common ground of what we categorized as technical studies (Gilbert et al. 2008) is that similar to case-based studies, most papers do not articulate a clear connection between HRODM investment and organizational effectiveness. Moreover, they provide limited scientific evidence to aid decision-makers concerning whether to adopt organizational data mining.

4 HRODM: Practical Implementation Tools and Expected ROI

4.1 Implications for Organizations

Our review of the literature underscores the importance of two notable fields within the HRODM research, namely, empirical and conceptual research. We further explore specific HRODM tasks and challenges in light of practical implementation tools and the expected ROI within organizational functions (Bassi 2011; Buede et al. 2018).

From a practical perspective, the ROI-based approach presented is important for a data-driven decision-making process in the field of HRODM. It also provides a step-by-step procedure for handling data and subsequently utilizing these data to attain meaningful managerial insights. Moreover, the need for a better focus in conducting and implementing HRODM projects within organizations is clear. Albeit with an element of shortage, some HRODM efforts in organizations today could be defined as reactive rather than proactive. Hence, it is not unusual for practitioners to use data that they receive access to in order to perform interesting analyses by addressing a question or set of questions with various levels of viability to the organization (Huselid 2015). The ROI-based approach to HRODM presented in this study provides a robust tool to compare and contrast different dilemmas and associated values that can be derived from conducting various types of ODM projects. The ROI-based approach also supports continuous improvement in organizations (Delbridge and Barton 2002).

From a theoretical perspective, the proposed categorization (Gilbert et al. 2008) provides a robust ROI framework for conducting research in the field of HRODM, thus enabling scholars and practitioners to focus on a desired topic in a more structured manner (Becker 2009; Lipkin 2015; Rasmussen and Ulrich 2015; Ghosh and Sengupta 2016; Pape 2016).

In Table 5, we illustrate the implications of HRODM for organizations. We present how addressing organizational challenges using various analytical tools,
Table 5  HRODM implications for organizations: Practical implementation tools & expected ROI

<table>
<thead>
<tr>
<th>Task</th>
<th>Sample challenges</th>
<th>Toola</th>
<th>Expected ROI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry Analysis</td>
<td>Macro market effect on turnover</td>
<td>Descriptive</td>
<td>Low</td>
</tr>
<tr>
<td>WORKFORCE PLANNING</td>
<td>High-demand jobs and attrition Person-Organization Fit</td>
<td>Predictive</td>
<td>High</td>
</tr>
<tr>
<td>Job Analysis</td>
<td>Robustness of job components</td>
<td>Descriptive</td>
<td>Low</td>
</tr>
<tr>
<td>Recruitment and Selection</td>
<td>Person-Job Fit</td>
<td>Predictive</td>
<td>High</td>
</tr>
<tr>
<td>Training and Development</td>
<td>ROI in training</td>
<td>Descriptive and Predictive</td>
<td>Medium</td>
</tr>
<tr>
<td>Compensation</td>
<td>Total compensation scenarios</td>
<td>Descriptive and Predictive</td>
<td>Medium</td>
</tr>
<tr>
<td>Performance Management</td>
<td>Performance management cycle scenarios</td>
<td>Descriptive</td>
<td>Low</td>
</tr>
<tr>
<td>Retention</td>
<td>Can retention be predicted</td>
<td>Descriptive and Predictive</td>
<td>Medium</td>
</tr>
</tbody>
</table>

Based on: Chalutz Ben-Gal (2019)

*Descriptive analytics tools may include: Descriptive statistics, Graphs and plots, Benchmarking tools, KPIs-Based Methods (scorecards), Business Intelligence (BI) Dashboards, and Advanced Survey Analytics

*Predictive analytics tools may include: Market Basket Analysis, Regression and parametric modeling (including Logistic Regression), Time Series Analysis, Classification Methods (e.g. decision trees, SVM, Discriminant Analysis, Neural Networks, Deep Learning), Clustering (K nearest neighbors, K-means), Anomaly detection, Profiling, Association rules, Link-Analysis, Causality modeling (Bayesian networks), Text Analysis & NLP and Attrition Modelling

namely, descriptive and predictive, may impact the expected ROI. This analysis may further assist scholars and practitioners in the ongoing effort to improve HRODM tools and impacts (Rasmussen and Ulrich 2015; Chamorro-Premuzic et al. 2017; Strohmeier 2018).

Table 5 presents the implications of HRODM for organizations as well as practical implementation tools. Specifically, it offers a summary overview of eight major tasks and activities that organizations are faced with, including their corresponding sample challenges (Srinivasan and Chandwani 2014; Bamber et al. 2017), practical implementation tools, and the expected ROI. The expected ROI is categorized into three levels – low, medium, and high – in accordance with the complexity of data-handling procedures that are relevant to the HRODM research (Fitz-Enz 2009; Rasmussen and Ulrich 2015; Ghosh and Sengupta 2016). The results documented in Table 5 yield two notable conclusions. The two areas of tasks that yield the highest ROI are workforce planning and recruitment and selection because both rely on predictive analytics tools (Fitz-Enz 2009).

As presented in Table 5, the first task focuses on *industry analysis*. This task ensures the analysis of basic parameters in an organization’s specific industry (e.g., retail, financial, technology). Empirical research tools are descriptive analytics that use BI and benchmarking to analyze government data, consulting firms’ data, census data, and macro-industry data. Our observations suggest examining macro market...
effects on specific constructs, such as turnover. Relevant ratios include industry average job turnover, average cost per hire, and job-specific retention budget, among others. Accordingly, the ROI for performing an industry analysis utilizing these HRODM tools is expected to be low.

Workforce planning is the second task, and we call for extended empirical analytics on this task because we believe it entails a high ROI. Workforce planning ensures the use of a continual process to align the needs and priorities of the organization with those of its workforce to ensure that it can meet its legislative, regulatory, service and production requirements as well as long- and short-term organizational objectives (Huselid 2015). Empirical research tools include predictive analytics that use various analysis techniques based mostly on internal data (e.g., ERP, headcount, product mapping, financials, budget) and external data (e.g., surveys, salary tables, syllabuses, and training program materials). Our observations suggest that certain challenges to test are person-organization fit and the connection between high-demand jobs and attrition. Accordingly, the ROI for performing workforce planning using these HRODM tools is expected to be high.

The third task focuses on job analysis, which is a process to identify and determine in detail a given job’s duties, requirements, and interfaces as well as its relative importance. This is a process in which judgments are made about data collected for a job (Levenson 2005). Empirical research tools include descriptive analytical tools (e.g., financial ERP, organizational structure, and headcount). Our observations suggest that some specific challenges to test are the robustness of job components and their effect on satisfaction and retention. Accordingly, the ROI for performing job analysis using these HRODM tools is expected to be low.

Recruitment and selection of talent is the fourth task, and we call for extended empirical research on this task because we believe it entails a high ROI. Practical research tools include predictive analytics. Our observations suggest that specific challenges to test are methods of classifying the talent pool according to available organizational resources; text analysis of interviews and profiling of vacant roles and organizational requirements; and logistics regression or other parametric models that predict recruitment probability of success, satisfaction, and person-job fit. Accordingly, the ROI for the recruitment and selection of talent using these HRODM tools is expected to be high.

The fifth task refers to training and development, which is primarily concerned with organizational activity aimed at improving the performance of individuals and groups in the organization. The recommended empirical research tools include both descriptive and predictive analytics. Our observations suggest that some specific challenges to test are the analysis of training ROI (through BI), whereas classification methods may assist in improving the training investment per job class. Accordingly, the ROI for performing training and development using these HRODM tools is expected to be at a medium level.

The sixth task refers to compensation and benefits. This management challenge assists in the execution of organizational strategy and may be adjusted according to business needs, goals, and available resources. Empirical research tools are descriptive (e.g., BI, scorecards, or other KPI-based methods) and predictive analytics.
Our observations suggest that some specific challenges are total compensation scenario testing; Monte Carlo simulations assess various compensation plans and regression analyses and their interplay with selected organizational phenomena. Accordingly, the ROI for performing compensation research using these HRODM tools is expected to be at a medium level.

The seventh task refers to performance management. This task is an ongoing process of communication between a supervisor and an employee that occurs throughout the year in support of accomplishing the organization’s strategic objectives (Huselid 2015). Future empirical research tools are based on descriptive analytics. Our observations suggest that some specific challenges to explore are performance management cycle scenarios mainly through BI, dashboarding, and KPI-based methods. Additionally, various levels of performance are clustered for the purpose of performance evaluation. Accordingly, the ROI for performing performance management using these HRODM tools is expected to be low (Buede et al. 2018).

Finally, the eighth task that is illustrated in Table 5 is retention of talent. The recommended empirical research tools are based on descriptive and predictive analytics. We call for a specific challenge to test and believe that further research on the topic of whether retention can be predicted is required. This challenge can be addressed by the profiling of key jobs, the classification of various talent retention scenarios, logistic regression, anomaly detection, and attrition modeling for various job groups. Accordingly, the ROI of performing talent retention using the proposed HRODM tools is expected to be at a medium level.

Our observations documented above apply to both scholars and practitioners when planning their future HRODM priorities. Furthermore, the ROI-based approach, which is the focus of this study, underscores the call for a more systematic approach for researchers and decision-makers to use evidence-based information as a guide to the adoption of HRODM and to understand its effectiveness (Rasmussen and Ulrich 2015; Buede et al. 2018).

The challenge for future research in HRODM is to reach beyond general studies in order to identify important contextual variables of HRODM and to consistently add value to existing organizational systems on both the contextual and practical levels. As emphasized earlier, we believe that much of this potential added value lies within the empirical and conceptual research in HRODM. Therefore, fertile avenues for future research contributions should focus on both empirical and conceptual studies in HRODM since these are the noticeable directions where the highest return rates are expected. Enhancing and developing empirical and conceptual knowledge in HRODM and state-of-the-art tools may serve as adequate future contributions to the field of HRODM.

If decision-makers have ROI information to guide the adoption of HRODM, a more focused and systematic research approach must evolve. Macro-organizational theoretical frameworks can add to the ROI-based approach by proposing different perspectives. For example, the contextual approach (Johns 2006, 2018) may offer a basis for understanding the organizational context in which specific ROI is to be found in line with new scholarly insights in the HRODM field. Additionally,
further theory development may integrate the LAMP framework (Boudreau and Ramstad 2006) with contextual elements (Johns 2006, 2018), which may also offer an appealing framework for testable hypotheses. Future rigorously constructed research questions may focus on the various tasks of HR from a holistic point of view while challenging the recommended analytical approach presented in this paper. Finally, future research may propose a new methodology that differs from the ROI-based approach to systematically analyze the scholarly and practical field of HRODM.

5 Contributions: ROI – Model to Guide the Way Forward

This section answers the third research question, RQ3: What is the future of HRODM research? It also reiterates the study’s contributions and emphasizes the ROI approach as a model to guide the way forward in HRODM research.

HRODM is a fascinating, dynamic discipline (Levenson 2011; Huselid 2015). The dynamic role of ODM enables it to focus both on the operational tasks of HR and on the organizations’ long-term and strategic business objectives. The growing field of HRODM enables management and engineering scholars and executives to implement a broader approach, which may increase their strategic contribution (Kazakovs et al. 2015; Strohmeier 2018). Machine-learning and data analytics in general, and more specifically in the field of HRODM, can aid in making informed decisions based on knowledge extracted from the available data and options (Sharma et al. 2015).

Our unique synthesis of the literature underscores the importance of two important fields within HRODM, namely, the empirical and the conceptual research. Our observations that we analyze and discuss in this study offer an ROI-based perspective to the HRODM field. Moreover, the ROI-based approach on the topic of HRODM presented in this paper provides theoretical and practical contributions. As a result, it provides a model to guide the way forward in HRODM research. From a theoretical perspective, this paper assists data analytics scholars who may find the ROI-based framework useful when fine-tuning their theoretical contributions in the field. From a practical perspective, this paper clarifies the dilemma associated with the HRODM field and assists practitioners regarding the expected ROI of HRODM initiatives within their organizations.

In conducting our ROI-based review of the literature on HRODM by integrating the analysis above, several major conclusions emerge. First, there is a need for more scientific empirical research in HRODM. Focusing on the development of such research might increase the potential for action-oriented, data-driven research, which can assist management and technical professionals. Second, as with the previous conclusion, and in light of notable deficiencies in the existing HRODM literature (Boudreau and Ramstad 2006; Dastyar et al. 2017), there emerges a need to focus on an ROI-based approach, which is our proposed model to guide the way forward.
We have taken a step toward systematically explaining some notable questions in the HRODM field. Not only does a focus on an ROI-based approach improve the adoption of HRODM as an important field in data science, but the context in which it is being adopted and implemented also matters, both practically and theoretically speaking.

6 HRODM Implementation: Two Examples

In this section, we provide two examples of HRODM implementation. The first example presents a Workforce Analytics and Big Data Analysis of Employee Turnover example – taken from Sela and Chalutz Ben-Gal (2018). The second example presents an Employee Recruitment Decision-Making Support Tool taken from Pessach et al. (2020).

6.1 HRODM Implementation Example I: Workforce Analytics and Big Data Analysis of Employee Turnover

Workforce analytics and turnover are strategic domains in HRODM implementation (Chalutz Ben-Gal 2019; Hausknecht 2014). This example shows that by utilizing HRODM tools for the purpose of workforce analytics through big data analysis and cluster analysis, management can gain a nuanced perspective on employee turnover and career path trajectories. This knowledge is important for strategic workforce planning across various industries. Furthermore, the digital world we live in results in weaker social connections between individuals. This drives the development of new workforce models, which are based on social networks and artificial intelligence (Fire and Puzis 2016; Aghabaghery et al. 2020). For example, in recent years, labor markets experience fundamental changes (e.g., social networks platforms and artificial intelligence-based recruitment and placement processes). Additionally, more candidates utilize online-based tools in their job search (Sela and Chalutz Ben-Gal 2018).

In this numerical example, taken from Sela and Chalutz Ben-Gal (2018), the authors extracted a unique dataset of over 970,000 curriculum vitae (CVs) based on LinkedIn profiles. Their analysis revealed just a slightly opposite relation between employee job satisfaction and turnover rate. The researchers found that while higher compensation packages provided by companies often lead to higher employee job satisfaction, they do not ensure lower turnover rates (Hausknecht 2014). The authors demonstrate that by utilizing HRODM tools, surprising and non-intuitive patterns are revealed, especially in global high-technology companies (see Fig. 3a, b).

In their research, the authors implemented the following methodology. They calculated the average employment period based on a dataset of 973,134 CV’s
retrieved from active LinkedIn profiles. Their dataset included 44 features, including control variables. For example, gender, country of employment, seniority, endorsed skills, as well as employment archival data – company name, job title, employment duration in previous firms, industry sector, etc. The dataset was merged with two additional benchmark data sources (Fortune 100; Glassdoor). The authors present their analysis and findings in Fig. 3.

Figure 3a presents machine-learning-based big data analysis of employee turnover. Figure 3a presents the average employee job satisfaction level, the average

![Job Satisfaction, Average Work Duration and Salary](image)

**Fig. 3 (a–d)** HRODM implementation: Example I – workforce analytics and big data analysis of employee turnover. (a) Big data analysis of employee turnover. Average employee job satisfaction (x-axis), average work duration in months (y-axis) and average salary in 39 organizations. Organizations such as Facebook or Google, have high salaries, high employee job satisfaction level, however a low average employment duration. In comparison, Intel has an average level of employee job satisfaction, but longer working durations. Symantec and Bristol-Myers Squibb both offer lower salary levels, relatively higher employee job satisfaction levels, and longer employment durations. Finally, Apple and EMC have rather low employee job satisfaction levels, high-average salary levels, and average work durations. (b) Analysis of employee turnover. Comparison of employment (work) duration histograms for eight companies: Facebook, Google, eBay, IBM, Apple, 3M, Intel, and Motorola (ordered by descending average employment periods). (c) Workforce Analytics of Career Paths. Three career path network clusters: financial cluster (red); consulting cluster (green); and the high technology cluster (blue). (d) Workforce analytics of career paths. Analysis of career paths within and across career network clusters reveals that Facebook and Google dominate two distinct employment clusters; IBM is a “Career Hub”. (Taken from: Sela and Chalutz Ben-Gal (2018))
Fig. 3 (continued)
employment (work) duration, and the average salary in thirty-nine organizations. As demonstrated by the authors in Fig. 3a, the companies Google and Facebook are located in the bottom left corner. These companies offer higher compensation packages (the red color indicates a compensation package of 130–140 k USD per year), and demonstrate higher job satisfaction levels. Utilizing HRODM tools,
the authors are able to demonstrate some surprising findings. For example, the researcher’s results indicate that Facebook has an average employment duration of 16.9 months. As extracted from their research and illustrated in Fig. 3a, organizations such as Facebook or Google, which offer high compensation packages, which seem to influence higher employee job satisfaction levels, demonstrate a surprisingly low employment periods for their employees.

The researchers compare and contrast their findings to other organizations. For example, they found that Intel, which demonstrates an average employee job satisfaction score, shows a relatively very long employment duration. They also found that, both Symantec and Bristol-Myers Squibb have lower compensation packages, however demonstrate higher job satisfaction levels, as well as longer employment durations. Finally, the researchers found that Apple or EMC have low job satisfaction scores, high compensation packages, and demonstrate average employment durations.

HRODM tools enable, researchers and practitioners alike, further data-driven analysis of the workforce. Furthermore, HRODM tools enable a granular view of the data. In Fig. 3b Sela and Chalutz Ben-Gal (2018) present a comparison of eight company’s employment (work) duration histograms. In Fig. 3b, the authors present histograms representing employment duration for eight technology companies. For example, while the upper three figures (representing Facebook, eBay, and Google) resemble an exponential function, the lower two histograms (representing Intel and Motorola) have a peak at an employment period of approximately 24 months, and smaller bins in lower and in higher values. The authors conclude that assuming an entry-level job, which usually requires about 6 months up to 1 year in order to reach mastery level of performance, these patterns are unexpected compared to Facebook or Google, where the average employment period is 16.9 and 23.3 months, respectively.

The authors emphasize the counterintuitive nature of their findings and claim that as Fig. 3a, b indicate, despite higher levels of employee job satisfaction and higher compensation packages, both Facebook and Google demonstrate shorter employment periods.

However, it is management’s role to analyze, interpret, and ultimately act upon such results, thus maximizing organizational ROI (Chalutz Ben-Gal 2019). Therefore, HRODM tools may assist the understanding of such macro market phenomena. For example, technology companies such as Google, Facebook, and eBay may be perceived as technological trendsetters, thus serve as “career platforms” for candidates toward their next desired job or professional challenge (Sela and Chalutz Ben-Gal 2018).

HRODM tools enable a deeper understanding of the workforce flows by analyzing employees’ career trajectory patterns. The researchers further illustrate an HRODM implementation example across industries. In Fig. 3c, the authors present three main career network clusters extracted from the LinkedIn dataset: the financial cluster (represented in red color), the consulting cluster (represented in green color), and the high technology cluster (represented in blue color).
The cluster analysis methodology presented by the authors, a popular tool within the HRODM domain, enables to detect employment and career moves for the purpose of workforce analytics. The authors claim that examining Fig. 3c, it seems that employees tend to make more frequent career moves within their own career network cluster. For example, employees working in the financial cluster (represented in red color), tend to make career moves to other financial companies, but less frequently to other career network clusters. Similar career trajectory patterns exist in the consulting cluster (represented in green color), and the high technology cluster (represented in blue color). Moreover, within the consulting cluster, the authors identified IBM as a “career hub,” which they define as a company that serves as a central crossroad junction for employees from which they can easily transfer to a different career network cluster. The author’s finding is surprising when compared to their Facebook and Google results, because both companies’ positioning represent a less central point as potential employers, and thus do not serve as industry hubs. Consequently, the authors conclude that one can detect that working at IBM may serve as a strategic career bridge to other attractive employment industries, compared to working in other companies, in which employees are more likely to stay within the borders of their career network cluster, especially when performing career choices (Sela and Chalutz Ben-Gal 2018).

Additional findings are illustrated by the authors in Fig. 3d, in which it is shown that both Facebook and Google are companies that dominate two distinct employment clusters (i.e., industries). Applying HRODM tools, the researchers analyzed employment clusters within a network that consists of almost 50 thousand individual career moves. They found that only one single career move was performed between these two companies. They also found that IBM is centrally located within its employment cluster, based on its physical central location, as well as on its evident large size.

This example emphasizes HRODM utility in the integration of machine-learning and data-driven tools, in order to maximize organizational ROI (Sela and Chalutz Ben-Gal 2018; Chalutz Ben-Gal 2019; Aghabaghery et al. 2020).

6.2 **HRODM Implementation Example II: Employee Recruitment Decision-Making Support Tool**

Employee recruitment is a strategic domain for HRODM implementation and is associated with high organizational ROI (Chalutz Ben-Gal 2019; Pessach et al. 2020), because it improves fit levels (Johns 2006, 2018) between a candidate and a specific position to be staffed. Recruitment and selection of talent is an organizational task associated with predictive analytics tools. Some HRODM application domains include methods of the talent pool classification; text analysis of interviews and profiling of vacant jobs compared to organizational demand;
prediction models for recruitment probability of success. Accordingly, the ROI for the recruitment and selection of talent using HRODM tools is expected to be high (Vihari and Rao 2013; Chalutz Ben-Gal 2019).

In this numerical example, taken from Pessach et al. (2020), the authors extracted a unique dataset and illustrated an application of a decision-making support tool for organizations and for the Human Resources community in order to improve the accuracy of recruitment and placement decisions. The example utilizes HRODM-driven machine-learning models for the prediction of recruitment success, as well as for extracting interpretable insights.

The authors measure the recruitment success based on a combination of the candidate turnover rate (Hausknecht 2014; Sela and Chalutz Ben-Gal 2018), and an objective target indicator based on performance. The authors also measure the performance indicator based on the position-changing conditions. For the purpose of classification and prediction of successful and unsuccessful recruitments and placements, as well as for mining significant patterns, the researchers use a Variable-Order Bayesian Networks Model (VOBN) (Ben-Gal et al. 2005; Singer and Ben-Gal 2007).

The authors evaluate the model compared to other machine-learning models applied to an extracted unique organizational dataset. The dataset utilized by the researchers includes about 700,000 cases of employee candidates who were recruited to an organization throughout a period of a decade (hired between the years 2000 and 2010). The authors detail some pre-hire features in the dataset. For example, position requirement, age, gender, marital status, education, grades, skills, interview and test scores, professional preferences, and additional socio-demographic features. Furthermore, the authors describe pre-processing activities. For example, data tables consolidation, sensitive data masking, etc.

In line with HRODM tools and techniques, the authors identified several clusters of position groups. Furthermore, using statistical data extracted from the Central Bureau of Statistics enabled the researchers to enrich the dataset with additional socio-demographic features. Missing values were tagged by zeros, and candidates with many missing values were removed by the researchers. In line with HRODM practices, additional dimensionality reduction procedures were performed by the researchers in accordance with the applied machine-learning algorithms.

The authors classified a “successful” and “unsuccessful” recruitment process as follows: They analyzed key reasons for employee turnover and categorized them into two groups: “successful recruitment” group, i.e., turnover associated with “natural” reasons, such as promotion (Hausknecht 2014), and an “unsuccessful recruitment” group, i.e., unexpected turnover, such as short-term or poor performance based turnover. Additionally, turnover was classified by the researchers as negative (e.g., “misfit”) or positive (e.g. “promotion”). Finally, the combination of turnover and position changes was utilized as a combined measure for labeling “successful” vs. “unsuccessful” recruitment.4

4 In line with HRODM practices, and in order to maintain consistency, the researchers applied an a-priori distribution of the target class on both a training and a testing datasets.
In this example, the authors were able to prove that the VOBN Model performs well in terms of both interpretability and accuracy in predicting recruitment success because it identifies context-based patterns that can support the organization in the recruitment process. Therefore, the authors explain that it can be used to extract rules and actions for the recruiters who sometimes lack the HRODM and machine-learning background, providing actionable and implementable insights – See Fig. 4a, b below. This HRODM implementation example is clear and easy to understand, therefore allows for an examination HR policies and procedures by what the authors call “extraction of interpretable and actionable insights” (Pessach et al. 2020).

In Fig. 4a – taken from Pessach et al. (2020) – the authors present the predicted probabilities of success in assigning sixteen candidates of two types to four positions by the machine-learning models (e.g., VOBN Model). For clarity purpose, the authors present colored entries in order to differentiate between the success probability values (i.e., high probability marked in green and low probability marked in red).

![Figure 4](image-url)

**Fig. 4 (a, b)** HRODM implementation: Example II – employee recruitment decision-making support tool. (a) Predicted probabilities of success in assigning sixteen candidates of two types of populations to four positions. The entries are color-coded by the success probability values. High probability (green), low probability (red). (b) Assignment of candidates to positions by four different solutions. Solution 1 (red) suggests the following: (i) recruiting 4 candidates to position 1409; (ii) recruiting 6 candidates to position 1509; (iii) recruiting 6 candidates to position 379 (note that none of them are of type 1); and (iv) not recruiting any of the candidates to position 40. (Taken from: Pessach et al. (2020))
The data presented by the researchers in Fig. 4a includes four positions (columns), sixteen candidates (rows), and two types of candidates who need to be assigned in a pre-determined way (e.g., based on their specific background, skills, and departmental demand). Figure 4a also presents the predicted success probability for each pair of candidate and position (the authors use shades of green to represent high probability and shades of red to represent low probabilities). For calculation purpose, the authors assumed a demand of six employees per position. Analyzing their dataset under these constraints, the authors conclude that based on the machine-learning algorithms if the goal is to maximize the sum of recruitment success probabilities, then position 379 requires staffing by type 2 candidates only.

In Fig. 4b, the researchers present four solutions to the allocation problem based on four different formulations where the rows represent candidates and the columns represent positions.

This example illustrates how HRODM techniques can be implemented as an employee recruitment decision-making support tool. This tool can support managers and recruiters alike when seeking candidates to be placed in target vacant positions. Moreover, the proposed decision-making tool illustrated in this example can further
assist the HR function in making relevant strategic decision. For example, employee development and retention procedures, in order to maximize organizational ROI (Chalutz Ben-Gal 2019; Pessach et al. 2020).

References


Singh, N. K., & Roushan, R. K. S. (2013), “Data Analytics and Decision Support in the context of ERP and beyond for giving CSIR a Competitive advantage” – two case studies from HR (Human resources) and MM (Materials Management) Modules.


