UNIVERSITY OF VAASA
SCHOOL OF MANAGEMENT

Anni Kaivola

CURRENT AND FUTURE TRENDS IN DATA DRIVEN TALENT IDENTIFICATION IN MNCS

Master's Thesis
in International Business

VAASA 2018

TABLE OF CONTENTS **page**

LIST OF FIGURES 5
LIST OF TABLES 5
ABSTRACT 7

1. INTRODUCTION 9
1.1. Research gap 11
1.2. Purpose of the study and research questions 11
1.2.1. Research question 12
1.2.2. Research objectives 12
1.3. Key concepts 12
1.4. Structure of the thesis 13

2. TALENT IDENTIFICATION IN MNCS 14
2.1. Definition 14
2.2. The ability component 15
2.3. The affective component 17
2.4. Talent identification process in MNCs 20

3. BIG DATA AND HR ANALYTICS 26
3.1. What is big data 26
3.2. Making sense of big data 28
3.3. Data analytics and HR 30
3.3.1. HR data 31
3.3.2. HR metrics 32
3.4. Data mining tools for personnel selection 33
3.4.1. Decision trees 33
3.4.2. Artificial neural networks 35
3.4.3. Support vector machines 36
3.5. Implementing HR analytics in MNCs 38

4. METHODOLOGY 40
4.1. Research approach 40
4.2. Research design 40
4.3. Data collection and sample 41
4.4. Data quality 44

5. FINDINGS **46**
5.1. Defining and identifying talent in MNCs 46
5.2. HR analytics and data in MNCs 48
5.2.1. Challenges related to data 49
5.2.2. Perceived benefits 50
5.2.3. Perceived challenges 51
5.2.4. HR ability to take on the data challenge 52
5.2.5. Visions for the future 53

6. DISCUSSION **55**
6.1. Defining and identifying talent in MNCs 55
6.2. Data, analytics and HR 59
6.2.1. The challenges 60
6.2.2. The benefits and future applications 62

7. CONCLUSIONS **65**
7.1. Managerial implications 66
7.2. Limitations and suggestions for future research 67

LIST OF REFERENCES **68**

APPENDIX 1. **77**

LIST OF FIGURES

Figure 1. Performance-potential –matrix 19
Figure 2. The potential pyramid 19
Figure 3. Talent multiplication model 23
Figure 4. LAMP framework 25
Figure 5. The machine learning cycle 29
Figure 6. Hyperplane in a non-linearly classifiable set 37
Figure 7. The HCM:21 model 38

LIST OF TABLES

Table 1. Skills, competencies and personality traits for talent identification 21
Table 2a. MNCs participating in the study 42
Table 2b. Consulting firms participating in the study 42
Table 3. Overview of interviews 42

UNIVERSITY OF VAASA
School of Management
Author: Anni Kaivola
Topic of the thesis: Current and future trends in data driven talent identification in MNCs
Name of the Supervisor: Jennie Sumelius
Degree: Master of Science in Economics and Business Administration
Department: Department of Management
Master's Programme: International Business
Year of Entering the University: 2016
Year of Completing the Thesis: 2018 **Pages:** 77

ABSTRACT

In today's data-driven business world information is omnipresent and companies are required to make smarter and faster decisions derived from data. At the same time, having the right talent, at the right time, in the right place is a challenge that companies continue to struggle with. Combining the data-driven decision-making is only starting to spread to the HR functions of global companies. Technological advancements are shaking HR functions and reshaping the way companies approach people management.

This master's thesis focuses on the intersection of talent identification and big data analytics tools where HR analytics and evidence-based talent identification are found. The purpose is to examine how big data tools can be used to facilitate talent identification in MNCs. Data and tools applicable for identifying talent are examined as well as the concept of big data and the tools to process it, such as artificial intelligence and its applications. Data mining techniques show great potential for more efficient and accurate personnel selection.

The empirical evidence in this qualitative, exploratory study was collected from eight MNCs and four consulting firms, through interviews with fourteen participants from Europe and North America. The findings of the study are discussed in connection to existing literature.

The findings indicate that HR analytics has not yet reached its full potential in the case organisations, and that further implementation is still underway. The HR function struggles to cope with the big data challenge. Predictive analytics are seen as a future trend that hold an immense amount of potential for increasing organisational performance from the people management perspective.

KEYWORDS: talent identification, HR analytics

1. INTRODUCTION

Today, more than ever before, there is an abundance of data and information available for organisations and managers. There are a great number of means to collect data as well as to store it, but what remains challenging is using that data in an efficient way. The abundance of available data and methods for processing it can facilitate managers' decision-making process in multinational companies (MNC) but at the same time it creates its challenges. (van Knippenberg, Dahlander, Haas & George 2015.) Proper analytical methods to make sense of the data can serve as a tool to facilitate decision-making processes and improve performance (McAfee & Brynjolfsson 2012; Sivarajah, Kamal, Irani & Weerakkody 2017), and organisations are engaging with human resource (HR) data and analytics. However, the challenge remains on how big data analytics can be used to create, capture, leverage and protect value from HR data, and how the existing HR analytics programs can be further developed to better measure and model the strategic impact of human capital inputs. (Angrave, Charlwood, Kirkpatrick, Lawrence & Stuart 2016.)

Talent identification can on its own be a challenging process for an organisation (Nijs, Gallardo-Gallardo, Dries & Sels 2014) and finding the right talent, at the right time, in the right place, is a real challenge in today's business environment where organisations face a diversified workforce (Beechler & Woodward 2009). Research shows that HR professionals often lack the skills, knowledge and insight to be able to ask the right questions when it comes to HR data they have at their disposal. (Angrave et al. 2016). This means that organisations are not using the available HR data to its full potential. HR has trouble structuring the data, as well as interpreting it correctly, and often managers make decisions based on data even though they do not fully understand it (Shah, Horne & Capellá 2012). The right questions enable accurate measurement of the valid metrics for HR analytics purposes, something that could and should be used in talent identification.

The role of HR has been evolving over the past decades from an administrative and maintenance-oriented function towards becoming more of a core business function and a strategic partner. Traditionally, HR has had an inside/outside perspective where its objectives have been to support the business, contribute to business results and add value. However, for HR to be able to add value to business and strategy also in the future, a shift towards outside/inside approach is needed. This means that the value is no longer created by only serving employees or redesigning HR practices for increased efficiency, but rather that value is created by aligning the internal HR services with the expectations from outside the organisation. Such approach can increase not only the value experienced by the organisation, but also the value for external stakeholders. One way of achieving this is by creating HR analytics focusing on the right issues and applying the appropriate HR skills to understand how the metrics can be used to support decision-making in the organisation and thus create value for both internal and external stakeholders. (Ulrich & Dulebohn 2015.)

HR analytics is not a new concept – the notion of measurement in HR can in fact be traced back to the beginning of the 20th century, and the first book on the modern era of HRM measurement was published in 1984. However, interest in HR analytics has increased strongly over the past few years (Marler & Boudreau 2017), and evidence-based HR and HR analytics have become more significant with the alignment of HR with business (Ulrich & Dulebohn 2015). Evidence-based management is about "making decisions through the conscientious, explicit, and judicious use of the best available evidence from multiple sources". This is done by taking a practical issue and turning it into an answerable question, systematically searching for and retrieving evidence, and critically judging, pulling together, and incorporating that evidence into the decision-making process, and evaluating the outcomes of the taken decision. (Marler & Boudreau 2017.)

According to Stahl, Björkman, Farndale, Morris, Paauwe, Stiles, Trevor and Wright (2012) the main idea of talent identification is to gain a better idea of what kind of human capital is working in the organisation and to get the personnel better aligned with the organisation's strategies. Talent identification is an important part of talent management and allows organisations to assess their human resources in a more defined and structured manner. In exclusive talent management, not everyone can be identified as talent, nor should they be since in organisations different types of employees are needed for different positions. The rationale underlying the exclusive view of talent management is that those who have been identified as talent are managed in a different way than those who have not – this way the organisation can get the most out of its employees. (Lepak & Snell 2002.) By aligning internal talent to a location where it is best suited and needed the most, and most importantly where it can meet not only current but also future challenges and opportunities, organisations access the full potential of their workforce. Through this type of categorisation and relocation of talent within the organisation the company can leverage its talent in the most effective way by allowing them to discover new capabilities within themselves and gain broader insight in the company's business from different business units. (Cheese, Thomas & Craig 2008: 78.)

1.1. Research gap

Even though HR analytics is an increasingly popular topic of discussion in the business world, it has not yet received much attention from academic scholars (Marler & Boudreau 2017). The adoption of HR analytics systems by organisations has been growing steadily over the last years, and a study shows that on average 39% of organisations in the US are using HR analytics to some extent. When focusing only on MNCs, 48% of organisations have some type of HR analytics system in use. (Harris & Spencer 2016: 70.)

Talent identification on the other hand has been researched extensively (see e.g. Spreitzer, McCall & Mahoney 1997; Pepermans, Vloeberghs & Perkisas 2003; Cheese et al. 2008; Tansley 2011; Dries & Pepermans 2008, 2012; Nijs et al. 2014). Having a clear vision of what is considered as talent in an organisation is essential for the task (Wiblen, Dery & Grant 2012). However, it is argued that there is little research on the practices of how organisations define and identify talent (Iles, Preece & Chuai 2010). Moreover, the intersection of talent identification and HR analytics has not yet been the focus of much research. Wiblen et al. (2012) conducted a study on the role of technology in talent identification and came to the conclusion that managers still rely on intuition even when there is technology in place aiming to facilitate the process. They propose that talent analytics, metrics and technology are key capabilities for future HR professionals in organisations, but discovered that the managers identifying talent in the case organisation lacked ability to interpret the relevant data and metrics, thus leaving them disregarded. Moreover, Aral, Bryonjolfsson and Wu (2012) focused on performance pay, HR analytics and information technology and discovered that organisations who adopted HR technology reached higher productivity than those who did not. Marler & Boudreau (2017) found Aral et al.'s (2012) study laudable and rare and suggest that similar approach could be applied to future research, where a well-accepted concept of HR – such as talent management – is examined with the additional effects of HRM information technology and HR analytics.

1.2. Purpose of the study and research questions

Many organisations are utilising big data and analytics for increased performance in a number of business functions, but as yet only few seem to utilise it for talent management purposes. (Harris & Spencer 2016:70, 76.) The idea of HR analytics is appealing to organisations but at the same time they struggle with its implementation – finding the right questions, the right metrics, and the right skills to analyse data to make decisions based on it. (Wiblen et al. 2012.) Thus, the purpose of the current study is to examine to what extent organisations create, capture and leverage value from available data through HR analytics to benefit the talent identification process.

1.2.1. Research question

The study focuses on the intersection of the talent identification process and the use of HR analytics in organisations. The overarching research question is formulated as follows:

> *How can big data analytics and tools be used in connection with talent identification in MNCs?*

1.2.2. Research objectives

To answer the research question, three objectives are set to indicate what the study is aiming to accomplish. The research objectives are the following:

> (1) *to understand how talent is defined in MNCs*
> (2) *to identify data analytics tools that are applicable for talent identification*
> (3) *to understand how analytics are used in MNCs to benefit the talent identification process*

1.3. Key concepts

In this section the following three key concepts of this study are defined: *talent identification, data,* and *HR analytics*.

Talent identification: Talent identification is the process during which employees' skills, capabilities, characteristics and other factors are assessed to predict talent within individuals (Nijs et al. 2014). This study focuses on internal talent identification, the maini idea of which is to gain a better idea of what kind of human capital is working in the organisation and to get the personnel better aligned with the company's strategies (Stahl et al. 2012).

Data: Data is the source of knowledge – it exists prior to argumentation or interpretation through which it is converted into facts, evidence and information. Data can be qualitative or quantitative; structured, semi-structured or unstructured; captured, derived, exhaust or transient; primary, secondary or tertiary; and indexical, attribute or metadata. (Kitchin 2014: 3–5.)

HR Analytics: HR analytics means the systematic identification and quantification of the people drivers of business outcomes. (van den Heuvel & Bondarouk 2016). HR analytics can be seen as an HR practice that is enabled by information technology. It uses descriptive, visual, and statistical analyses of data and big data related to HR processes, human capital, performance, and external economic benchmarks to create business impact and enable evidence-based, data-driven decision-making. (Marler & Boudreau 2017.) The purpose of HR analytics is to enable the making of better decisions concerning HRM in organisations. Terms such as workforce analytics and people analytics coexist and can be used interchangeably with the term HR analytics. (van den Heuvel & Bondarouk 2016.)

1.4. Structure of the thesis

The first chapter of the thesis is the Introduction where the topic is introduced, and the research gap justified. The objectives of the study are set, and the key concepts are briefly defined. The literature review of the study is divided into two main chapters. Chapter two explores what is meant by talent, its different components and how talent is identified in MNCs. Chapter three introduces big data and data analytics, and the concepts of artificial intelligence, machine learning and data mining. HR analytics are also discussed as well as the intersection of talent identification and HR analytics. For this, data analytics and HR are explored, and the most relevant data mining techniques are presented and discussed in connection to personnel selection. Chapter four introduces the methodological choices of the study and presents the data. The findings of the empirical study are presented in chapter five. Chapter six discusses the findings of the study in connection with the literature review. Finally, chapter seven concludes with a summary of the findings and a discussion of the managerial implications and the limitations of the study. Also, suggestions for future research are provided.

2. TALENT IDENTIFICATION IN MNCS

This chapter provides a more detailed discussion and definition of talent. The different components of talent and defining characteristics, competencies and personality traits of talent are examined. Moreover, the talent identification process is discussed.

2.1. Definition

The main idea of exclusive global talent management and identification is to gain a better understanding of what kind of human capital is working in the organisation, to get the personnel better aligned with the company's strategy, to retain and engage the organisation's key employees and to address the need for global managerial competencies (Stahl, Björkman, Farndale, Morris, Paauwe, Stiles, Trevor & Wright 2012, Scullion & Collings 2011: 7–9). Talent management is something that organisations often struggle with and Collings and Mellahi (2009) argue that it is due to the lack of consistent definition and clear conceptual boundaries across the field. In their view, the starting point for all talent management in organisations should be defining talent. They suggest the following definition of talent management: "activities and processes that involve the systematic identification of key positions which clearly contribute to the organization's sustainable competitive advantage, the development of a talent pool of high potential and high performing incumbents to fill these roles, and the development of a differentiated human resource architecture to facilitate filing these positions with competent incumbents and to ensure their continued commitment to the organization." (Colligns & Mellahi 2009: 2.) This definition suggests that the talent management practices should be rooted in the organisation's strategy.

Much of the research is focused on the talent, the "A players" (eg. Tansley 2011; Lepak & Snell 2002; Nijs et al. 2014), while some academics choose to adapt an approach where the strategic key roles are identified before looking more closely at the workforce (eg. Collings & Mellahi 2009; Huselid, Beatty & Becker 2005; Boudreau & Ramstad 2005). These "A positions" are defined by their significant importance and direct impact to the organisation's ability to execute some part of its strategy, and their high-performance variability among the employees in the position (Huselid et al. 2005). Only after the strategic key roles in the organisation have been identified should the discussion on who is talent begin. Many researchers have tried to put into words the essence of *talent*. In short, individuals classified as talent by the organisation are considered to have the potential to reach high levels of performance (Tansley 2011), make a positive difference in the organisational performance (Chartered Institute of Personnel and Development 2007), and their human capital rates high on both uniqueness and strategic value (Lepak & Snell 2002). Most organisations associate talent with individuals who show the most potential to move up the organisational ladder to senior roles and leadership positions (Tansley 2011). Moreover, Silzer and Church (2010) identified six different approaches to defining talent in organisations. First, by role, refers to the potential to move into top or senior management roles. Second, by level, refers to the ability to perform two levels above the current role. Third, by breadth, refers to the capability to take on broader scope with more demanding tasks in a leadership role. Fourth, by record, refers to a consistent track record of exceptional performance. Fifth, by strategic position, refers to the key positions that are at the core of organisational success and sixth, by strategic area, refers to functions, business units or geographical regions which are central to the organisation's strategy. (Silzer & Church 2010.)

The purpose of talent identification should not only be to recognise the talent already expressed within the organisation, but also to identify the employees who have the potential to excel in different and larger roles in the future which is why performance should not be used as a sole measurement for talent but also ability and affective components should be considered while determining talent in the organisation. The former refers to the innate abilities in a specific domain of human functioning and systematic development, while the latter refers to motivation to invest and interest. Even though performance is an important part of talent definition, it is not the most crucial. As suggested, the ability and affective components of the individual should be emphasized over performance results. (Nijs et al. 2014.) We are first going to look more closely at the ability component, followed by the affective component.

2.2. The ability component

Various researchers have adopted different approaches when determining the characteristics and predicting qualities of *talent*. There are different perspectives to talent, such as behavioural aspects, knowledge, skills, and competencies and cognitive ability (Tansley 2011). There is also discussion on whether talent is innate, or if it can be developed later in life. When talent is seen as innate, the conversation is built around stable competences such as ability and personality. On the other hand, if talent is viewed as something that can be developed later in life the focus is on dynamic competences such as knowledge and skills, which can be acquired through training. It is also suggested that the two types of competences are intertwined in a way that stable competences facilitate the acquiring and development of dynamic competences. (Leiba-O'Sullivan 1999.) A competency is defined as a capability or ability which is a set of different but related sets of behaviour organised around an underlying construct which Boyatzis (2008) calls intent, which many academics have referred to as meta-competence (eg. Le Deist & Winterton 2005; Briscoe & Hall 1999; Brown & McCartney 1995). Meta-competences are defined as follows: "meta-competences are the higher-order skills and abilities upon which competences are based and which have to do with being able to learn, adapt, anticipate and create, rather than with being able to demonstrate that one has the ability to do. [---] [M]eta-competences are those abilities, skills, and capacities which exist above and beyond any competence which an individual may develop, guiding and sustaining them, and from which they originate." (Brown & McCartney 1995: 47–48). Stable competences and meta-competences are practically synonyms for each other as meta-competences are described to be the facilitators of acquiring other substantive competences (Le Deist & Winterton 2005; Brown & McCartney 1995; Boyatzis 1982, 2008).

When determining the competences that are characteristic to talented individuals, there are different approaches. Differentiating between competences and meta-competences is done in a similar manner as the differentiation between stable and dynamic competences. Le Deist and Winterton (2005) suggest a holistic model of competence in which the dimensions of cognitive competence, social competence and functional competence are bound together by meta-competence as the input that facilitates the acquisition of the output competence. A model based on three clusters of competencies which differentiate talent from the workforce can be used when seeking to identify talent. The three clusters are cognitive competencies, emotional intelligence competencies and social intelligence competencies. It is suggested that competencies are a behavioural manifestation of emotional, social and cognitive intelligence. (Boyatzis 2008.) Also, the different perspectives of talent – behavioural aspects, knowledge, skills, and competencies and cognitive ability – further aid in identifying talent in the organisation. Behavioural aspects include resilience and confidence – a can-do attitude based on self-belief. Knowledge and skills refer to having enough creative ability to construct new realities and experiences and create new knowledge. Competencies and cognitive ability refer to the diversity of thought or flexibility in creating a state of mind that matches the organisation's requirements and are irrelevant in terms of a particular job role. (Tansley 2011.)

While the range is wide with a myriad of definitions and specified vocabulary, looking more closely into the long lists of competencies from different academics, the similarities are evident. For example, Dries and Pepermans (2012) categorise skills and personality traits to analytical skills, learning agility, drive and emerging leadership while Tansley (2011) identifies the broad categories of abilities, aspirations and engagement. Analytical skills mean intellectual curiosity, strategic insight, decision making and problem solving. Learning agility refers to willingness to learn, emotional intelligence and adaptability. Drive means result orientation, perseverance and dedication. Emerging leadership refers to motivation to lead, self-promotion and stakeholder sensitivity. (Dries & Pepermans 2012.) Comparing the categorisation of Dries and Pepermans' (2012) to that of Tansley's (2011) there is commonness that can be seen. The abilities include innate characteristics, mental and cognitive ability, emotional intelligence, learned skills and technical skills. Aspirations refer to the extent to which an individual desires prestige and recognition in the organisation, advancement and influence. Finally, engagement means the willingness to go above and beyond the average job demand, and the passion and motivation of the employee. (Tansley 2011.) All of the analytical skills and most of the learning abilities identified by Dries and Pepermans (2012) fall under Tansley's (2011) identified abilities. The emerging leadership traits can be found in both abilities and aspiration while all the personality traits in the category of drive are found under the components of engagement.

2.3. The affective component

The ability component of talent emphasises the intellectual abilities of individuals whereas the affective component considers the non-intellectual attributes and their effect on individual performance. These non-intellectual attributes include personality, interests, and motives. (Nijs et al. 2014.) There is much research on personality, and the common consensus is that personality consists of five relatively independent dimensions. These five dimensions are known as the Big Five – extraversion, emotional stability, agreeableness, conscientiousness, and openness to experience. (Barrick & Mount 1991; Leiba-O'Sullivan 1999.) Moreover, there are lists of personality traits that are associated with the five different personality dimensions. For example, extraversion relates to traits such as assertive, talkative, active initiative, ambition and expressive. Emotional stability refers to feelings of anxiousness, depression, anger, embarrassment and insecurity. Agreeableness is associated with traits such as courteousness, flexibility, good-natured, cooperative and tolerant. Conscientiousness relates to dependable, thorough, responsible, hard-working, achievement-oriented and persevering personality traits. Finally, openness to experience refers to traits such as imaginativeness, curiosity, broad-minded and intelligence. (Barrick & Mount 1991.)

While personality alone does not predict talent, the personality traits associated with the Big Five relate to the competencies and meta-competency discussed in the previous section. Also, personality and competencies together do not predict talent either. Some researchers suggest disregarding lists of competencies, skills and personality traits all together and focusing on measuring the required outcomes of each employee (Buckingham & Vosburgh 2001). However, this is not the popular approach among academics as majority of talent management literature posits that talent identification should not be based uniquely on performance (eg. Nijs et al. 2014; Tansley 2011; Wiblen, et al 2012; Dries & Pepermans 2012). An important part of talent identification is the measurement of potential (Tyler 2011; Silzer & Church 2009a).

Potential can be understood as the possibility of an individual becoming something more than what they currently are and implies further growth and development in order to reach a desired end state. The term *potential* suggests that an individual has a set of qualities which enable him/her to succeed in different or larger roles in the organisation in the future and is thus more associated with the future possibilities than with current performance. (Silzer & Church 2009b.) A widely used tool for categorising employees is the performance-potential –matrix or '9-box grid' which can be seen in Figure 1. (Tyler 2011.) There is some criticism of the tool found in literature as it is argued to create confusion between potential and performance (Robinson, Fetters, Riester & Bracco 2009). Some suggest that it is not adept for measuring potential (Lumme-Tuomala 2017), which is true as its purpose is to categorise employees rather than rate them (Tyler 2011). Due to the confusion of the two concepts, potential and performance, Robinson et al. (2009) presented an alternative method for assessing potential in the organisation. Their suggestion is a model where performance is treated as a component of potential and it is shaped as a pyramid. The model is an ascending hierarchy of decision steps where the employee is examined at each checkpoint to see if they meet or exceed the criteria to move on to the next step. The organisation's mission, value and culture are at the base of the pyramid, followed by consideration of the employee's performance in their current role and his or her behaviour. Meeting the criteria means moving up in the pyramid. However, if the criteria are not met, the model provides suggestions regarding the employee's role in the company. This can be seen in Figure 2.

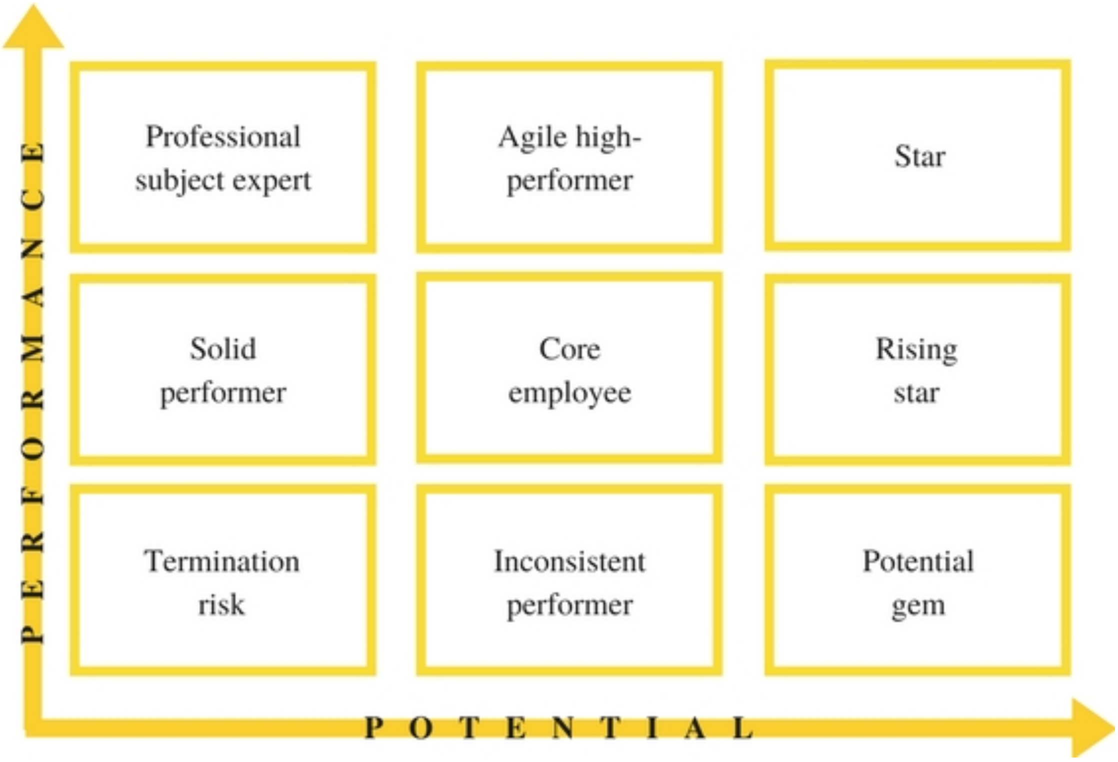


Figure 1. Performance-potential –matrix, also known as the 9-box grid (Tyler 2011)

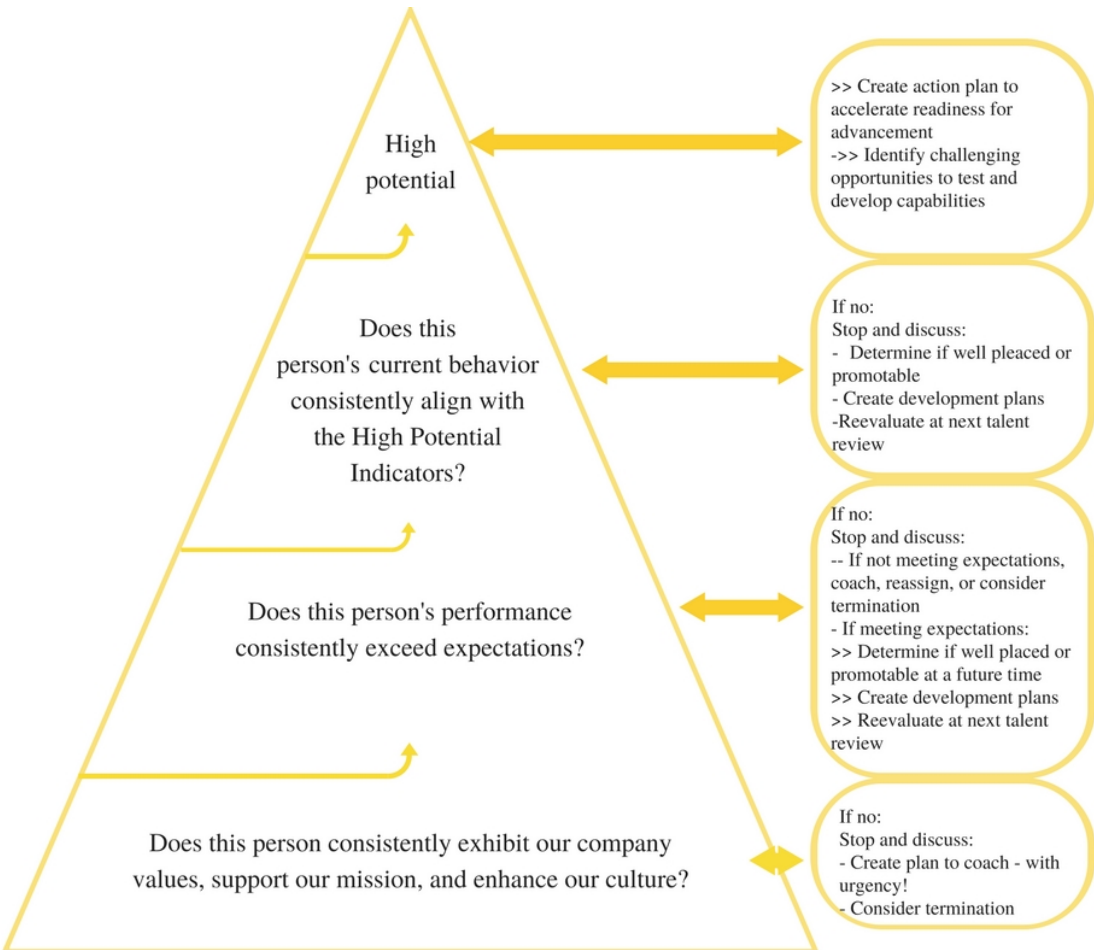


Figure 2. The potential pyramid (Robinson et al. 2009: 414)

The problem with both of these models is their emphasis on performance and the ability component of talent (Lumme-Tuomala 2017), and they disregard to an extent the two remaining dimensions of the affective component of talent – motivation to invest and interest (Nijs et al. 2014). The affective component of talent relates to the concept of potential which can be determined by a person's ability, engagement and aspiration to rise and succeed in more senior and critical positions (Lumme-Tuomala 2017). Aspiration translates to the extent to which a person wants or desires prestige and recognition in the organisation, advancement and influence, financial rewards, work-life balance, and overall job enjoyment (Tansley 2011, Lumme-Tuomala 2017). A great indicator for potential in an employee is their commitment to excel in the pursuit of unselfish goals. Employees with potential are ambitious while aspiring to big, collective goals. They show deep personal humility and invest in getting better at everything they do. Alongside with motivation, four other predicting qualities of potential are curiosity, insight, engagement and determination. Curios people seek out new experiences, knowledge and candid feedback as they are marked by an openness for learning and change. Insight refers to the ability to gather and make sense of information in novel ways. Engagement means the tenancy to use emotion and logic in communicating a convincing vision and connecting with others. Finally, determination dedication to fight for achieving even difficult goals despite challenges and to bounce back from adversity. (Fernández-Aráoz 2014.)

In order to better capture and summarise both the ability and affective components of talent, the most significant skills, competencies and personality traits identified in previous research as commonly considered for identifying talent are summarised in Table 2.

2.4. Talent identification process in MNCs

Talent identification is a process which remains hierarchical in many organisations as the initiative and lead is taken by the immediate supervisor or the top management. Allowing fully open initiative from anyone in the organisation concerning the identification process is rare. (Pepermans et al. 2003.) Talent identification is commonly done by board members and managers, albeit in different combinations. During the process employees' different skills, capabilities, characteristics and other factors are assessed to predict talent within individuals. (Dries & Pepermans 2008, Nijs et al. 2014.) Large organisations such as MNCs are likely to formalise their talent identification process to ensure no one in the organisation is not overlooked and that the process is

Table 1. Skills, competencies and personality traits for talent identification

Meta-competence (Brown & McCartney 1995, Le Deist & Winterton 2005)	**Competency**	**Related skills**	**Related personality traits**	**Authors**
Managing, functional competence	Leadership	Visionary thinking, influence, persuasion, people development, economic value creation	Aspiration, motivation to invest, interest	BCG (2008: 50), Tansley (2011), Nijs et al. (2014), Brown & McCartney (1995)
Cognitive competence	Learning ability, cognitive ability	Analytical skills, diversity of thought	Flexibility, willingness to learn	Tansley (2011)
	Analytical skills	Decision making and problem solving, pattern recognition	Result orientation, perseverance, risk taking	Dries & Pepermans (2012), Spreitzer et al. (1997), Brown & McCartney (1995)
Influencing, social competence	Emotional intelligence	Teamwork, cultural knowledge, interpersonal skills, relationship building	Empathy, sensitivity	Tansley (2011), Dries & Pepermans (2012), Pepermans et al. (2003), Brown & McCartney (1995)
	Technical skills	Ability to generate revenue and customer satisfaction,		Wiblen et al. (2012)
Personal competence, social competence	Behavioural aspects	Seeking and using feedback, self-control	Resilience, confidence, cogency, curiosity, initiative, tenacity, flexibility, drive	Tansley (2011), Pepermans et al. (2003), Spreitzer et al. (1997)

carried out within deadlines (Silzer & Church 2010). To better organise the talent within the organisation, most organisations use talent pools. Talent pool is a collective of employees who have been identified as talent by the organisation. (Tansley 2011.) Talent pools can be gathered for various purposes. For example, managerial talent, functional level talent, all-around managerial talent from different levels of the organisation, or fast-track programs for graduates. This shows that talent in organisations is segmented both vertically and horizontally, which is essential in managing the workforce to it best performance. (Dries & Pepermans 2008.)

Talent identification is done in order to better leverage the human capital working in the organisation (Stahl et al. 2012) and it is a key part of talent management. By assessing the abilities of the human resources, employees can be categorised and thus managed more efficiently. Those who have been identified as talent are often managed in a different way than those not labelled as talent, and by doing so the organisation can optimise its human resources to their maximum potential. (Lepak & Snell 2002.) Achieving this means aligning the internal talent to a location where it is best suited, most needed, and where it can meet not only the current but also the future challenges and opportunities (Cheese et al. 2008: 78). Organisations should identify their need for talent before anything else, which will allow more accurate and relevant measurements to base talent decisions on (Nijs et al. 2014).

Despite the benefits of talent identification and talent management, MNCs continue to struggle with their practices. Inadequate talent management practices in organisations can in the worst-case lead to failure in the ability to accurately identify talent. (Cappelli 2008, Nijs et al. 2014.) Research stands to show that systematic identification of talent in an organisation will lead to a high organisational performance (Collings & Mellahi 2009), yet organisations use inconsistent and ad hoc methods for identifying and developing their talent (McDonnel, Lamare, Gunnigle & Lavelle 2010). Assessing the whole workforce in all its components will allow the organisation to better manage its employees towards excellent performance in the future by finding and discovering activities that better fit their motivation and interest areas. (Nijs et al. 2014.) A complete assessment of the workforce allows the organisation to assign tasks more effectively. By categorising the workforce into different pools of candidates, it is easier to put together teams for projects and thus maximise the potential of the employees for the best organisational performance. (Collings & Mellahi 2009; Tansley 2011.) The starting point for talent identification should be deciding on the high-potential categories to consider, the talent and potential definitions, and on which organisational levels the talent discussion will be applied to (Silzer & Church 2010).

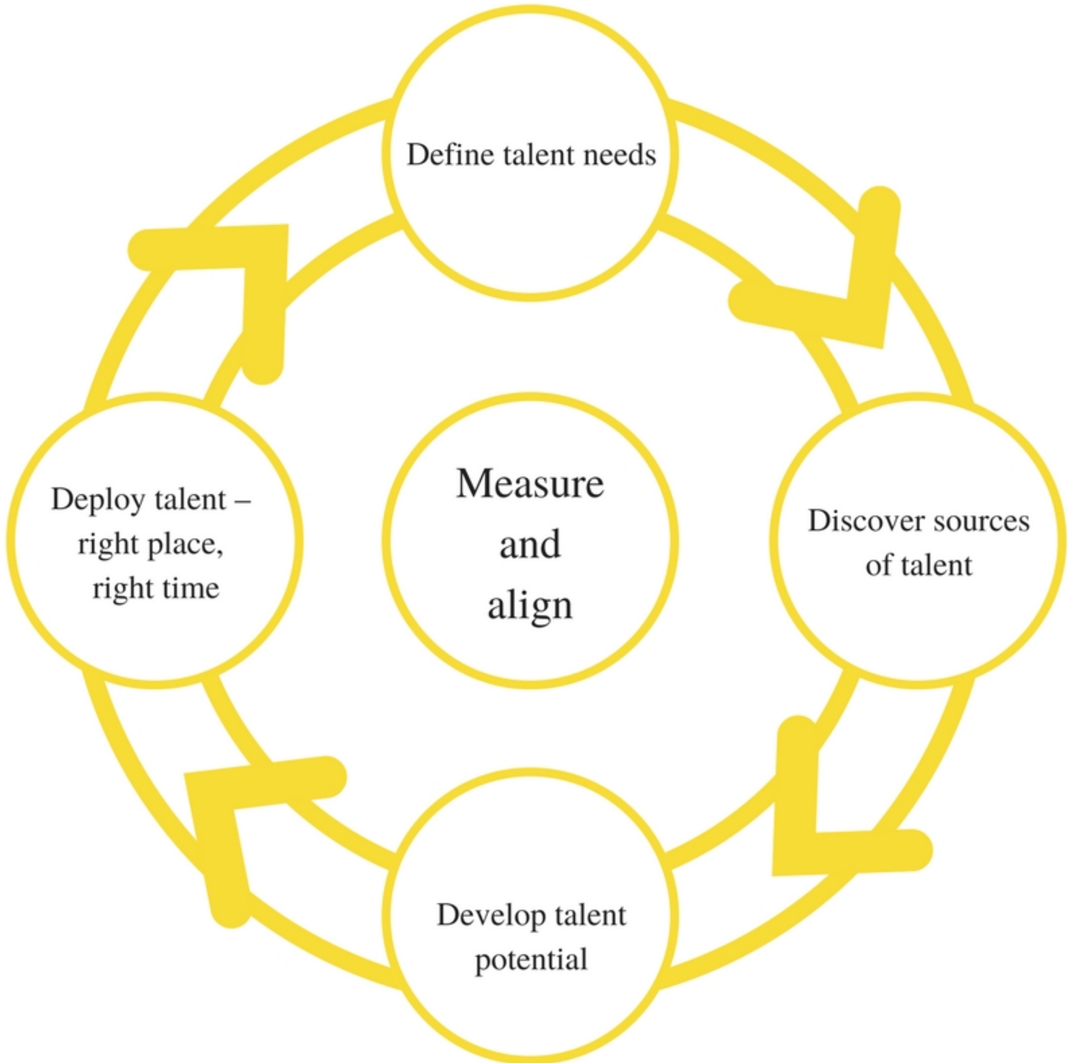


Figure 3. Talent multiplication model (Cheese et al. 2008: 85)

Talent powered companies are high-performing organisations and deploying talent in the right ways will create a synergy effect of talent multiplication. Identifying talent will allow the organisation to cultivate their talent and to combine it in different ways for superior levels of effort, imagination, creativity, learning, adaptability, and performance from the entire organisation. The talent multiplication model starts by defining the organisation's talent needs and moves on to discover the sources of talent. After this, the talent potential is developed before the final stage of deploying the talent in the right place at the right time. These four stages form a continuous cycle throughout which measurement and alignment are performed. (Cheese et al. 2008: 57–58.) This is visualised in Figure 3. For the model to work effectively in an organisation, the mindset needs to move from personnel control and people development towards the integration and alignment of talent management activities and the business strategy (Cheese et al. 2008: 85). The model provides a framework for consistent talent management practices yet implementing analytics into the talent management process might need more effort.

Cascio and Boudreau (2008) have developed a framework for evidence-based talent identification – the LAMP. The letters stand for *logic, analytics, measures*, and *process*. These are critical components of measurement system that can affect strategic change and organisational effectiveness. Embedding HR measures within the LAMP framework, organisations can create change through enhanced decision-making. (Cascio & Boudreau 2008:8.) The logic component of the framework emphasises the thought processes for identifying the most critical questions. Sufficiently deep logical framework will articulate the specific connection point between HR investments, as well as their effects on talent pools and organisational outcomes. The potential measurement and analysis are more precise with precise logical questions. (Boudreau & Ramstad 2006; Cascio & Boudreau 2008: 9–10.) Analytics in turn "transform HR data and measures into rigorous and relevant insights" (Boudreau & Ramstad 2006). Thus, analytics are an essential addition to logic in order to build an effective measurement system (Boudreau & Ramstad 2006; Cascio & Boudreau 2008: 11–12). Having relevant measures for the organisation's purposes will provide more valuable analytics outputs. There is no sense in measuring something that does not have strategic value for the development of the organisation. The process component refers to having sufficient management practices in place for learning and knowledge transfer. (Cascio & Boudreau 2008: 10–13.) The LAMP framework is illustrated in Figure 4.

Figure 4. LAMP framework (Cascio & Boudreau 2008: 8)

3. BIG DATA AND HR ANALYTICS

First, this chapter examines big data before moving on to discuss how it can be treated. Concepts of artificial intelligence, machine learning and data mining are introduced and the connection between data analytics and HR is discussed. Further, three data mining techniques and their applicability to HR analytics are explored.

3.1. What is big data

In the digitalised world of information technology, organisations have access to a large amount of data and companies sometimes appear to collect data just for the sake of it. The abundance of information available along with the possibilities for processing all the data can facilitate decision-making processes for managers in organisations, but at the same time it remains challenging to use the data in an efficient way. (van Knippenberg et al. 2015.) Big data are defined by five key characteristics: volume, velocity and variety. First, big data are huge in *volume* as the data consists of terabytes or petabytes of data. The second character, high in *velocity*, refers to the data being created in or near real-time. Finally, the data are diverse they are structured and unstructured in nature and often referenced temporally and spatially. There are other defining characteristics for big data beyond the 3Vs, such as *exhaustive, resolution, indexical, relational*, *flexible* and *scalable*. Big data being exhaustive refers to the scope of striving to capture entire populations or systems. When talking about resolution it is meant that big data are fine-grained and aim to be as detailed as possible and uniquely indexical in identification. The relational nature of big data means the data contains common field which enable the conjoining of different datasets. Big data are extensional as in new fields can be added easily, and thus described by flexibility, and big data can be expanded in size rapidly, making the data scalable. (Kitchin 2014: 68.) As big data are usually drawn from multiple sources that are beyond the control of a single actor, it is difficult to handle the data in traditional ways (Janssen, van der Voort & Wahyudi 2017).

There are different types of challenges related to big data depending on the life cycle of the data. The first challenges faced are about the data itself—its volume, velocity, variety, variability, veracity, visualisation and value. The next challenges are about the process challenges which are an issue when processing the data. For example, data aggregation and integration, analysis and modelling, data interpretation, data mining and cleansing, or data acquisition and warehousing. (Sivarajah et al. 2017.) Van Knippenberg et al.

(2015) argue that data are accessible and the methods for cumulating and storing data develop continuously. This is an important issue since large quantities of data require special storage systems so that they can be accessed with ease. The last challenges data create are the management challenges such as privacy, security, data governance, data and information sharing, cost/operational expenditures, data ownership, and ethical issues related to data usage (Sivarajah et al. 2017).

The challenge is not in finding and accumulating data but rather in having a long enough attention span to analyse that data and to make use of it (van Knippenberg et al. 2015). Data only have utility when meaning and value can be extracted from them. This means that it is not enough to generate and store data just for the sake of having data – rather it is what is done with the data is that matters. (Kitchin 2014: 100.) Time is needed to analyse data as well as attention which are both scarce resources. More information scales faster than the attention of decision makers who have to decide which information has priority and what is disregarded. This essentially means that managers have to make a decision on what data to use and for what purpose before they can actually make the decision they set out to make. This kind of chain of thought takes up time and attention. Even without the abundance of data, people have a limited capacity to process information and to hold their attention, and to motivate themselves to acquire and absorb information. Thus, managers need to learn to identify the relevant information regarding the decision at hand meaning that the quality of attention is in key position. (van Knippenberg et al. 2015.)

Sivarajah et al. (2017) propose that big data can facilitate the decision-making process when the proper analytical methods are used to make sense of the data. Moreover, having efficient methods for analysing the data is essential. The development of highly powered computation has resulted in methods and tools for data analytics and approaches such as artificial intelligence and machine learning are ideally suited for handling and extracting information from large, connected datasets and big data. (Kitchin 2014: 100–101.) There are four basic approaches to analytics: description, explanation, prediction, and prescription. Descriptive analytics help understand what has happened by focusing on the data and information itself. The underlying questions are *what* and *when* did something happen and *how often* does it happen. Explanation analytics help understand *why* something has happened and what is its impact. Predictive analytics can help in creating insights on what could happen in the future through forecasts and statistical modelling. The underlying questions are *what is likely to happen next* and *what if we did this/that instead*. Prescriptive analytics can be used to assess how business can be enhanced on different levels while decreasing expenses by answering questions such as '*What is the optimal answer or outcome?*' and '*How is that achieved?*' (Kitchin 2014: 101, Sivarajah et al. 2017.) Additionally, Sivarajah et al. (2017) identify the concept of pre-emptive analytics which can help in taking precautionary actions on events that might have a negative effect on the organisational performance.

3.2. Making sense of big data

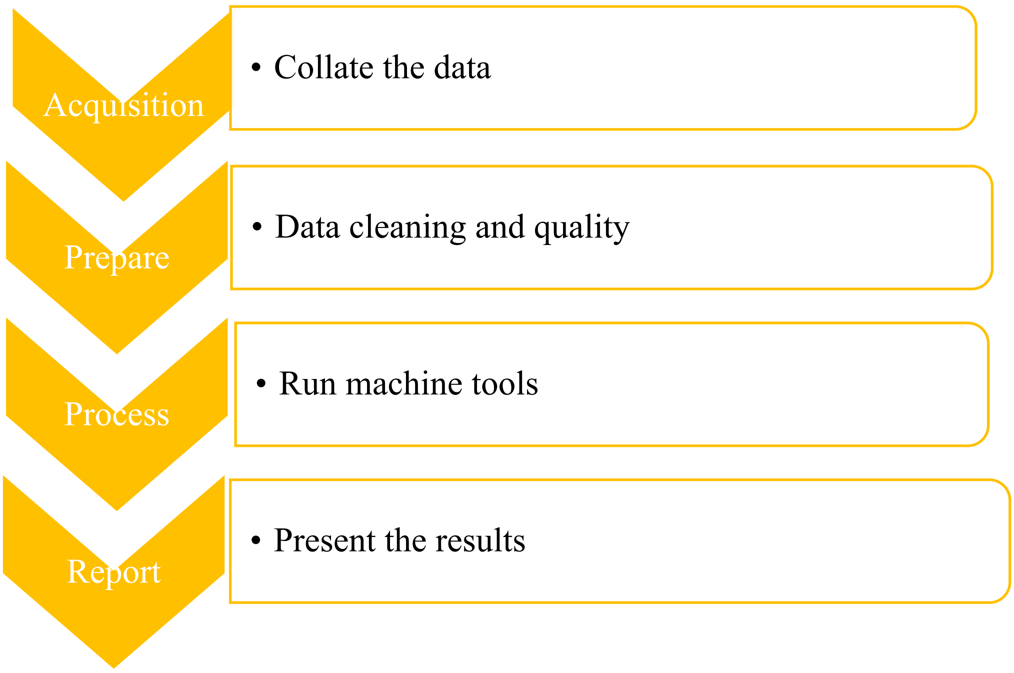
As the definition of big data suggests, it is challenging to make sense of large quantities of data (Sivarajah et al. 2017) and given the nature of big data – the volume, the velocity, and the variety – it is challenging and time consuming to approach the task without powerful computation and the help of artificial intelligence (Kitchin 2014: 100–101). Artificial intelligence as a discipline studies how to make computers perform tasks which, at the moment, people perform better (Ertel 2011: 2). The goal of artificial intelligence is to mimic human intelligence and cognitive ability which makes machine learning one of the most central subfields of artificial intelligence. Human intelligence is characterised by its ability to adapt and adjust to environmental conditions, and we are able to make changes in our behaviour accordingly through learning – an ability which in humans is vastly superior when compared to that of machines. (Ertel 2011: 1–3, Muhammad & Yan 2015.) Machine learning studies computer algorithms that improve automatically through experience (Ertel 2011: 164). A classic definition of machine learning is the following (Bell 2014: 2):

> "*A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with the experience E.*"

Machine learning seeks to evolve an understanding of a dataset through an iterative method. The purpose is to automatically learn to recognise complex patterns from the data and construct models which help explain and predict such patterns and optimise the outcomes. (Kitchin 2014: 103.)

Machine learning can be divided into two broad types of learning: supervised and unsupervised. The former refers to a situation where a model is trained with a training dataset to match inputs to certain known outputs. The latter refers to a situation where the model aims to teach itself to notice patterns and find structure in a large sample of data. This is done without the use of training data. (Kitchin 2014: 103, Bell 2014: 3.) Another way to describe unsupervised learning is that the model learns patterns in the input even when it is not given any explicit feedback (Russell & Norvig 2014: 705, Lim, Tucker & Kumara 2017). There is no right or wrong answer when building a model through unsupervised machine learning – the machine learning algorithm is run and the patterns, outcomes, clusters and relationships among data are observed (Kitchin 2014: 103, Bell 2014: 3, Russell & Norvig 2014: 705–706). Unsupervised clustering algorithms discover natural clusters without prior information (Lim et al. 2017). The creation of a machine learning algorithm begins by creating learning rules and weightings which shape the learning process and direct how the model is built in relation to the data. Building a model with machine learning begins with a simple construction which is tweaked repeatedly using the learning rules until a robust model is developed. (Kitchin 2014: 103.) The simplified machine learning cycle can be seen in Figure 5.

Figure 5. The machine learning cycle (Bell 2014: 17)



Machine learning is a powerful tool which can be employed to undertake different types of big data analytics (Kitchin 2014: 103). One of the most widely used approach to machine learning is data mining, which is "the process of acquiring knowledge from data, as well as its representation and application" (Ertel 2011: 166). Data mining is performed in large datasets and it is the process of extracting data and patterns. The underlying premise of data mining is the notion of all massive datasets holding meaningful information which in its nature is non-random, valid, novel, useful and ultimately understandable. (Kitchin 2014: 104.) Since data mining is a subcategory of machine learning, the process and logic behind are similar: a training dataset is used to learn and discover data patterns, which are then tested to see how well the made generalisations work (Ye 2014: 17). Depending on the type of data (structured, unstructured or semi-structured) and the purpose of the analysis, there are different methods of data mining to choose from (Kitchin 2014: 105).

3.3. Data analytics and HR

Predictive and descriptive analytics have been applied to many fields in organisations, such as finance, marketing and operations. However, applying analytics to the HR domain has been lagging behind in popularity and HR analytics are less evolved than their other domain counterparts. (Lismont, Vanthienen, Baesens & Lemahieu 2017.) Researchers are becoming increasingly interested in data analytics approaches most applicable to the HR function in organisations (eg. Jantan, Hamdan & Othman 2009abc, Chien & Chen 2008, Dursun & Karsak 2010, Lin 2010, Fan, Fan, Chan & Chang 2012). The most value can be created through Knowledge Discovery in Database (KDD) (Jantan et al. 2009c). KDD is a broad term which includes the concepts of machine learning, statistics, visualisation, database and expert systems (Adriaans & Zantinge 1996: 7) and data mining techniques, such as Decision Tree, Bayesian Fuzzy Logic, Support Vector Machine, Artificial Immune System, Neural Network, Rough Set Theory, Genetic Algorithm and Nearest Neighbour (Jantan et al. 2009c; 2010). There are commercial software, such as SAP SuccessFactors, OrgVue and Workday, which are HR business intelligence (BI) systems (SAP 2018, OrgVue 2018, Workday 2018). However, these HR BI systems are based on the paradigm that users know which questions need answers and what needs to be analysed. Such systems are built on relational databases where the user defines the relevance of a data item to answer a question. This decision is made before accessing the data. Big data systems on the other hand operate in large, diverse, and unstructured datasets where they make inferences about the probabilities of certain events or associations. Big data systems use machine learning and its applications to process the data. The machine is responsible for determining the relevance of data items based on statistical analysis – a decision which is made after accessing the data. (Pape 2016.)

Predictive analytics hold an immense amount of potential for organisational management and development. Above all, analytics are a mental framework and a logistical progression. Even though statistical operations have an important role, it is necessary to understand the interactions and the relationships of the elements for the problem at hand. (Fitz-Enz 2014: 2.) Analytics are a framework for logical progression and a set of statistical tools. Logic is used to create a design of the framework in which the statistical data is applied. The result is an evaluation of strategic, operational, and future opportunities. (Fitz-Enz 2010: 4–5.) The four basic approaches to analytics seen in 2.1. – description, explanation, prediction, and prescription (Kitchin 2014: 104, Sivarajah et al. 2017) – can be applied in the HR domain. However, the HR function suffers from a capability gap. While the general consensus among professionals is that said gap should be bridged by building and developing the appropriate analytical capabilities internally, there is little knowledge on how such in-house development should take place. (Minbaeva 2017.) Moreover, scholars are questioning the ability of HR function to effectively apply big data and analytics for organisational gain and suggest that HR does not have sufficient skills, knowledge and insight to ask the right questions (King 2016).

3.3.1. HR data

Before being able to apply any analytical approach to the domain of HR, data are needed. There are many types of data created in organisations. It is important to understand what type of data are being created and what type of data are needed in order to answer the questions that the used analytics aim to address. Information on, for example, finance accounts, client addresses and delivery status are data items which are essentials for the organisation's operations. Yet, there are much data that is not essential for running a business and are thus optional to create. (Pape 2016.) Many organisations choose to collect data on different aspects of the business, such as productivity and performance, in different departments of the organisation. However, the traditional silo structures of organisational functions are affecting the mentalities in data management: the data created in a department stays in that department. This kind of mindset prevents the HR related data from being combined with data from other departments in the organisation resulting in loss of valuable insights on the data. (Angrave et al. 2016.)

HR function creates a lot of data on employees. There is both numerical, quantitative data, as well as more verbal, qualitative data. The former are related to, for example, pay, hours worked, sales made, hours billed to clients, and other measures of individual output. The latter are related to, for example, performance appraisals, training and development experienced by the employee, grievances, internal communications, and staff attitudes surveys. (Angrave et al. 2016.) Traditional HR BI systems are equipped to work with such structured datasets (Pape 2016). However, more valuable insights from the data could be extracted when combined to 'bigger' data and leveraging big data systems instead (Angrave et al. 2016; Pape 2016). Such data can be related to what employees are doing, who do they communicate with, and what do they communicate about. This would mean location-based data from mobile devices, Internet browsing histories, email and phone records, data from online collaborative tools, email content, and instant messenger conversation logs. Such data could provide insight on mood and morale, social networks, and interactions within the organisations as well as with external stakeholder groups. Such analytics are challenging to implement as there are significant privacy issues related. (Angrave et al. 2016.) Privacy concerning big data is a significant issue. For instance, when personal information is combined with external large datasets, new information on the person is created which the person might want to keep private. (Katal, Wazid & Goudar 2016).

Such challenges affect the fact that data quality is one of the most significant barriers for adapting credible HR analytics in organisations (Minbaeva 2017). Also, it is suggested that the HR function lacks a detailed understanding of analytics, its approaches and its applications. This obstructs HR professionals from having meaningful interactions with data. However, the HR function is not the only one to blame as many analytics experts in turn do not understand HR. Such lack of overlapping skills, competences and expertise in the HR and analytics professionals results in a mismatch between what HRIS software can do and what HR departments need. (King 2016.)

3.3.2. HR metrics

HR metrics are used to analyse many things in organisations, such as the correlation between performance and compensation, predicting the number of employee turnover, the time to fill vacancies, employee count and estimated monetary value of performance difference in role. HR data items can be categorised, for example absence, application information, competences, contract information, cost of employee, disciplinary and grievances, employee communication, health and safety, well-being, hiring and induction, motivation, personal details, performance and potentials, role information, termination, and training. (Pape 2016.) HR metrics can also be categorised around performance in order to map the performance drivers, efficiency and performance itself. Examples of such could be, respectively, employee competency growth and extent to which employees are clear about their own goals, average time for dispute resolution and response time per information request, and competency development expense per employee and number of suggestions generated and/or implemented. (Becker, Huselid & Ulrich 2001: 64–71.)

3.4. Data mining tools for personnel selection

Talent identification is essentially a personnel selection problem and can be categorised as a classification problem (Chien & Chen 2008) which supports the choice of using data mining techniques such as decision trees, artificial neural networks, and support vector machines (Kitchin 2014: 104).

3.4.1. Decision trees

Decision trees aim to create a model that will predict the value of a target variable based on the set of input variables (Bell 2014: 45, Ye 2014: 37). Decision trees are a powerful learning algorithm for artificial intelligence as they are simple and efficient for extracting knowledge from data (Ertel 2011: 184). Also, it is a powerful tool for data mining, as the extracted knowledge can be easily understood, interpreted, and controlled by humans and reporting back to others is easy by simply referring back to previous steps in the tree (Ertel 2011: 184; Bell 2014: 46). Every tree is made of nodes which are associated with one of the input variables (Bell 2014: 48).

There are the following four elements in a decision tree: *decision nodes, chance nodes, utility nodes*, and the *edges* that connect the different nodes (Jensen 2001: 122, Bell 2014: 48). Decision tree always starts with a root node and ends on a leaf, which is a utility node. Moreover, the edges of the tree do not converge at any point. (Bell 2014: 48.) The edges between the nodes are labelled: an edge from a decision node is labelled with the chosen action, and an edge from a chance node is labelled by a state. Studying the decision tree starts from the root and continues downward. When passing a decision node, the label on the link indicates what the decision is. When passing a chance node, the label on the edge indicates the state of the node. Decision node following a chance node indicates that the node has been observed. There is an assumption of *no-forgetting* – after a decision has been taken, the decision maker know all the labels on the path from the root down to the current position in the decision tree. (Jensen 2001: 122.)

Decision trees are able to perform well even with large sets of data. However, depending on the data presented in the training set, the decision trees can create overly complex models. This can be addressed by reviewing the training data set and pruning values to categories. (Bell 2014: 46-47.) Identical subtrees within the decision tree can be collapsed. This creates a coalesced decision tree which is solved through the same procedure as normal decision trees. The purpose of this is to reduce complexity in the exponentially growing number of decision and chance variables. (Jensen 2001: 128.) Some commonly used algorithms for decision tree analysis are ID3 (Iterative Dichotomiser 3), C4.5, CHAID (Chi—squared Automatic Interaction Detection) and MARS (Multivariate Adaptive Regression Splines). (Bell 2014: 46-49.)

As decision trees are built on machine learning, they are able to analyse data without assumptions about the underlying distribution. (Ertel 2011: 184, Chien & Chen 2008; Jantan et al. 2010.) Chien and Chen (2008) propose a framework for exploring relationships between personnel profiles and work behaviour by applying the CHAID decision tree algorithm. Jantan et al. (2010) propose a framework using the C4.5 decision tree algorithm. Different decision tree algorithms have their advantages, and the accuracy of decision tree classifiers in talent forecasting has been studied. The study compared five different classifier algorithms: C4.5, NBTree, REPTree, BFTree and SimpleCart. The C4.5 classifier algorithm was found the most accurate with the rate of 95.14%, followed by SimpleCart (70.78%), BFTree (70.07%), NBTree (67.26%) and finally REPTree (65.76%). (Jantan et al. 2009b.)

Applying decision tree algorithms to personnel selection is built on the same framework as any decision tree. Chien and Chen (2008) identify six steps while Jantan et al. (2010) identify only three. However, the steps include very similar phases. First, the problem at hand needs to be defined and understood, and clear objectives need to be structured. Domain knowledge of personnel selection is very important at this stage. (Chien & Chen 2008.) Second, the data needs to be collected and prepared. It is important to understand what kind of data is relevant to include in the dataset with relation to the objective of the analysis. (Chien & Chen 2008; Jantan et al. 2010.) Third, the data mining model is construct. Depending on the type of data and the nature of the problem at hand, the decision between different algorithms is made. (Chien & Chen 2008.) After choosing the algorithm, classification rules are generated in a training dataset (Jantan et al. 2010). Fourth step is model analysis and evaluation, which is done to review the model. Fifth step is interpretation and knowledge extraction. Again, domain expertise is needed for justifying the meaning of the extracted knowledge. Further study can be initiated on unusual patterns or results in order to confirm their validity. (Chien & Chen 2008.) Final step is to use the discovered knowledge. The output of the algorithm can thus be used in a decision support system for personnel selection. (Chien & Chen 2008; Jantan et al. 2010.)

3.4.2. Artificial neural networks

Artificial neural networks are modelled on the parallel architecture of animal brains and it is based on a simplified input-output form. It is a computing system which is built from a number of simple and highly interconnected processing elements. These elements process information by a dynamic state response to external inputs. Artificial neural networks are used within real-time or near real-time scenarios as they excel on data volume and speed as they seek to mimic the brain with an increased speed factor. The neural networks are based on what is called the *perceptron*. The perceptron receives an input signal and passes the value through some form of function, the result of which is the output. The function that calculates the output is called the *activation function*. The result of the activation function determines if the value is passed to the output axon and to the next neuron in the network. There are many interconnected neurons in an artificial neural network, and they each have their own inputs, outputs, and activation functions. (Bell 2014: 91-97.)

Artificial neural networks solve problems in a nonlinear fashion and use multilayer perceptrons. This means that there are one or more layers between the input and output nodes. The layer or layers between the input and output layers are called *hidden layers.* The multilayer perceptron comes embedded with the concept of backward propagation of errors or *back propagation* in short. Back propagation has two steps: propagation and updating of the weight. Back propagation occurs in neurons of the network. As artificial neural networks are optimal for high volume and speed data, they are tolerant to noisy data. However, data preparation is still important because artificial neural network work only on numerical data values which means that data with text values need to be converted. (Bell 2014: 96-100.)

Even though the basic structure of artificial neural networks is simple, the model is capable of learning very complicated relationships, which supports its use in personnel selection (Huang, Huang, Huang & Jaw 2004; Wei & Xiaolin 2011). Evaluation of talent is based on multiple factors, as was seen in chapter two, which makes the decision-making environment very complex. Huang et al. (2004) suggest a model for identifying managerial talent in two dimensions: individual traits and managerial skills. The former include capability, motivation and personality, while the latter include conceptual skills, interpersonal skills, and technical skills. Zhang, Li, Li and Xiao (2013) conducted a study on evaluating the components of talent in information technology through a model based on artificial neural networks. Determining the number of layers, input nodes and output nodes should be based on the problem at hand. Domain expertise has a significant role in this. Further, it is suggested that using artificial neural networks can avoid human error in decision making due to their ability to self-organise, to self-adapt, and to make decisions in uncertain, paradoxical knowledge environments (Wei & Xiaolin 2011).

Goonawardene, Subashini, Boralessa and Premaratne (2010) test the accuracy of their proposed artificial neural network model for talent prediction with a manual process. The evaluation process revealed similar discovery to that of Wei and Xiaolin (2011): there is high uncertainty involved in the manual process which can be overcome with automated systems, such as artificial neural networks (Goonawardene et al. 2010). With profound understanding of the domain and problem at hand and with the appropriate input and output nodes, artificial neural networks are a powerful tool for talent prediction. The accuracy rates yielded by Zhang et al. (2013), Wei and Xiaolin (2011) and Goonawardene et al. (2010) were 93.3%, 92.31% and approximately 90%, respectively.

3.4.3. Support vector machines

A support vector machine is a technique for classifying objects (Bell 2014: 139). The approach brings together the advantages of linear and non-linear models. An algorithm is applied to non-linearly separable problems in a two-step process: 1) applying nonlinear transformation to the data with the property that the transformed data is linearly separable. 2) determining the support vectors in the transformed space. The central nonlinear transformation of the vector space is known as the *kernel*. (Ertel 2011: 252–253.) The support vector machine is a form of supervised learning which means that training data as well as test data are involved in developing the algorithm. As the approach is used for classifying objects, understanding binary and multiclass classification is essential. Binary classification is only applicable when there are two outcomes to the classification problem, the object is either one or the other. When there are multiple outcomes to the classification problem, it is called multiclass classification. (Bell 2014: 139–142.)

Linear classifiers are used to establish the location of the objects in order to determine the group an object belongs to. The line separating the groups is known as the *hyperplane*. The support vector machines define the maximum margin with the assumption that the hyperplane is separated in a linear fashion. Since the object are linearly separable, another two hyperplanes can be created. These hyperplanes are known as *edge hyperplanes* and they define the offset on both sides of the main hyperplane. There are no objects inside the area that spans between the main hyperplane and the edge hyperplanes. The objects on the edge hyperplanes are the support vectors. However, the linear model is not always applicable to the real world. This is why non-linear classification often has applications to real world problems. In cases where the objects stray from the hyperplane, an algorithm is applied to fit the hyperplane's maximum margin in a feature space. This function is the kernel function. (Bell 2014: 144-146.) The Figure 6. Illustrates the hyperplane in a set of objects that cannot be classified with a linear classifier.

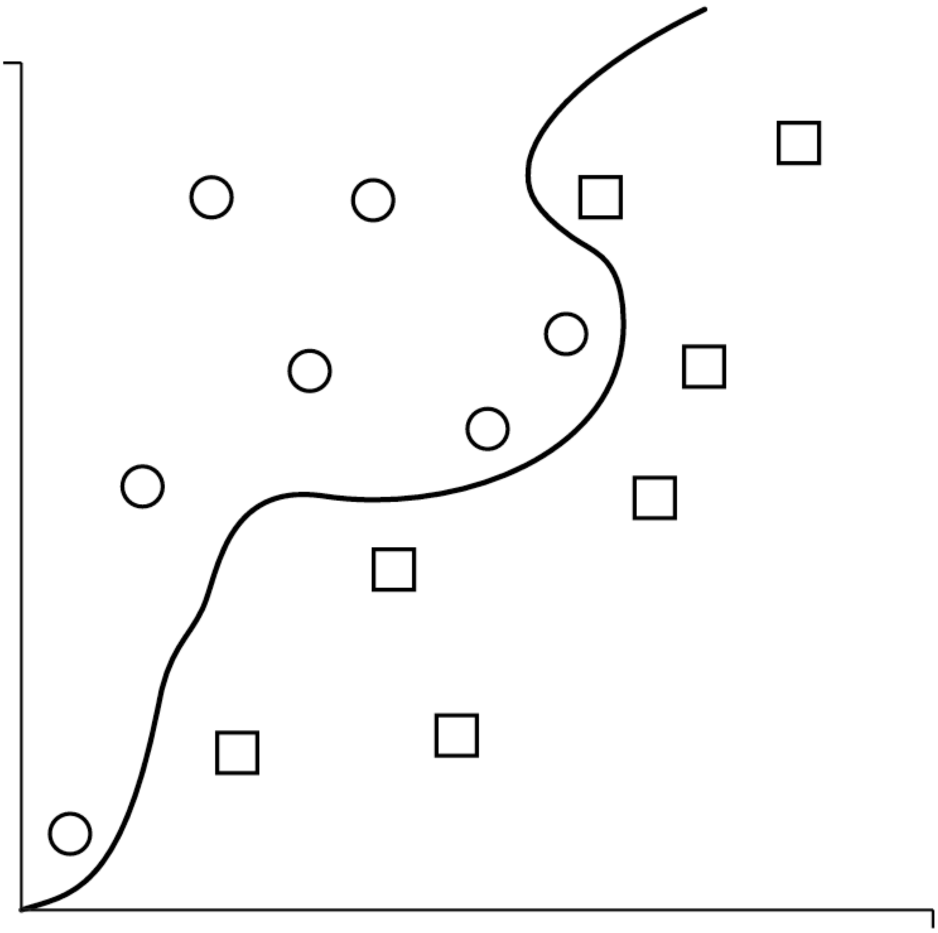


Figure 6. Hyperplane in a non-linearly classifiable set (Bell 2014: 147)

Support vector machines are found to be efficient in personnel selection as they are capable to model complex non-linear decision boundaries (Yasodha & Prakash 2012). Huang (2015) suggests using a least squares support vector machine approach for forecasting talent. The model uses an index system based on technological innovation talent stock competitiveness, talent development environment competitiveness, talent utilisation efficiency competitiveness and talent's sustainable development competitiveness. Yasodha and Prakash (2012) proposes a model using a class-attribute contingency coefficient (CACC) support vector machine classification algorithm for classifying HR data. The proposed model can be used to predict potential talent in organisation.

3.5. Implementing HR analytics in MNCs

Fitz-Enz (2010) introduces a model for human capital planning as opposed to traditional HR. The model is called human capital management for the twenty-first century (HCM:21) and it is a comprehensive approach that aligns different HR functions with the organisation's business vision, values, and plan. It offers a logical framework for gathering, organising, and interpreting data and knowledge for assessing probability of upcoming events. The model has four steps: scan, plan, predict and process. During the first step, both internal and external factors are assessed to develop an understanding of things that are likely to happen, what is the competition, and how do these affect the human, structural, and relational capital. (Fitz-Enz 2010: 5, 56, 86–87.) The HCM:21 model is illustrated in Figure 7.

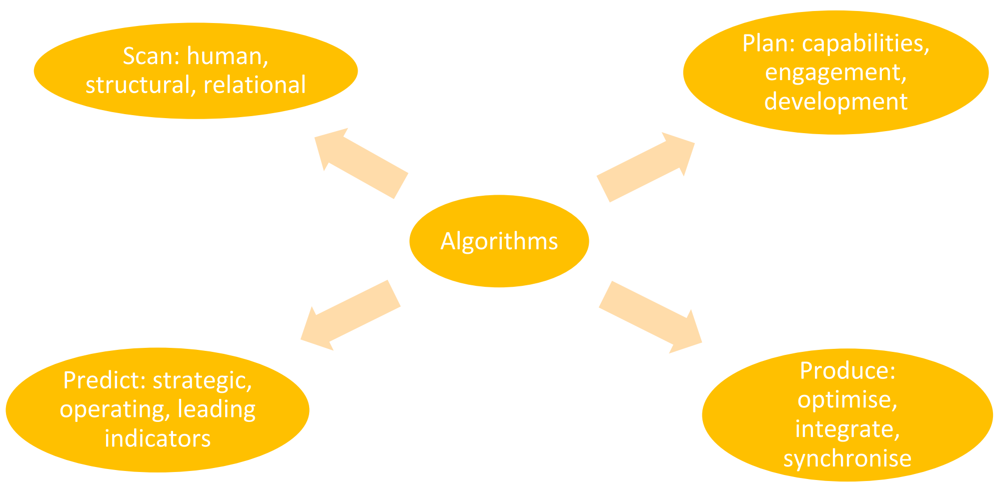


Figure 7. The HCM:21 model (Fitz-Enz 2010: 16)

The second step, plan, means reconstituting workforce planning as capability development to create an agile system focused on building human capability in the organisation through engagement. Third, produce, means optimising, integrating and synchronising the HR services into inputs, throughputs and outputs by applying statistical analysis. Finally, predicting refers to a measurement system that includes strategic, operational and leading indicators which are used to tell a story through causal and correlational aspects. (Fitz-Enz 2010: 6–7.) However, it is important to take into consideration the factors that might hinder the implementation of HR analytics in MNCs. The extent to which processes can be standardised across the company might pose challenges due to the institutional pillars that influence organisational practices – regulatory, cognitive, and normative. The legal and regulatory environment of a subsidiary can affect the data usage. The cognitive dimension affects interpretations and the frameworks establishing meaning. The normative dimension affects the values and norms in the society. Thus, it is likely that adapting HR analytics practices across MNC subsidiaries will pose challenges from standardisation perspective as the institutional environment varies from one country and region to another. (Heikkilä, Brewster & Mattila 2014.)

4. METHODOLOGY

This chapter explains the methodological choices of the study. First, the research approach and design are discussed. Then, the data collection is described, and an overview of the data sample and analysis is provided. Finally, data quality issues such as reliability, generalisability and validity are discussed.

4.1. Research approach

This study takes an abductive research approach (Saunders, Lewis & Thornhill 2012: 144–148) which is based on the pragmatist research philosophy (Wilson 2014: 8–11). Abduction refers to a process where both induction and deduction are used to understand the phenomenon described (Eriksson & Kovalainen 2016: 24) by moving back and forth between theory and data (Saunders et al. 2012: 147–148). In abduction, testable conclusions are generated from known premises. The purpose of data collection is to explore a phenomenon, and to identify themes and patterns. These are located in a conceptual framework drawn from existing theory. (Saunders et al. 2012: 144.) Pragmatists place centre focus on the research problem and the research questions and apply the methods they consider to be the most appropriate to generate significant insights on the topic (Wilson 2014: 10–11).

4.2. Research design

This study is exploratory in nature (Saunders et al. 2012: 171) and has a qualitative multiple-case study design which supports the research question: *how can big data analytics and tools be used in connection with talent identification in MNCs?* Exploratory studies are used to gain insights about issues, problems or phenomena (Saunders et al. 2012: 171) and case study strategy is commonly adopted when seeking to develop an understanding on why decisions were taken, how they were implemented and with what results (Yin 2009: 17). A case study strategy allows in-depth analysis of a phenomenon within its real-life context (Saunders et al. 2012: 179). Multiple-case study design is likely to produce more compelling evidence and has the capacity to demonstrate literal and/or theoretical replication (Yin 2009: 53–64, Saunders et al. 2012: 180–181). Case study research is well suited for the purpose as the form of research question is *how*, it does not require control over behavioural events and it focuses on contemporary events (Yin 2009: 8). Moreover, this study is cross-sectional, meaning that the data is gathered at one point in time and represents a snapshot of the phenomenon (Saunders et al. 2012: 190).

The primary data collection was conducted through a single data collection technique – semi-structured interviews. Semi-structured interviews are conducted with the help of a list of themes and possible key questions that need to be covered. The order of the questions may vary from one interview to another while some questions might be left out. (Saunders, Lewis & Thornhill 2016: 175, 184, 200, 391.) Semi-structured interviews are by nature non-standardised interviews; including face-to-face interviews, internet-mediated interviews via Skype, and group-interviews. Semi-structured interviews have a setting where the researcher has a list of themes and possible key questions to be covered, the use of which can vary from one interview to another. Semi-structured interviews are more frequently used for explanatory studies but can also be applied to exploratory studies. (Saunders et al. 2012: 374–375, 377.) Secondary data was drawn from online sources, such as the case company website and the websites of companies offering HR analytics systems and other tools for talent purposes. The use of both primary and secondary data – triangulation – is recommended for case study research as it increases the reliability and validity of the study (Yin 2009: 114–116).

4.3. Data collection and sample

The data collection for the study was conducted through semi-structured, qualitative interviews. Selected interviewees for the study were either representatives of an MNC or a consulting firm that provides solutions, tools or software for MNCs to take on the HR analytics challenge. A total of 13 interviews were conducted with 12 different companies – 8 MNCs and 4 consulting firms. See Tables 3a and 3b. The purpose of including both MNCs and consulting firms in the study is to gain wider perspective to what the current state of HR analytics is, and how the field is developing. A prerequisite for participation from the part of the MNCs was the use of HR analytics tools to some extent or their implementation in the near future. The prerequisite for participation from the part of the consulting firms was that they provide HR analytics solutions, tools or software to MNCs.

The data collection took place over a period of one month from the beginning of March 2018 to the beginning of April 2018. One of the MNCs was interviewed twice and one of the interviews was a group interview. The interviewees were given general themes of discussion prior to the interview. Based on the general themes of discussion the case companies themselves chose who in their organisation would have the most knowledge about the topic. The topics of discussion were (a) the company's view on talent, (b) tools, metrics and technology used for talent identification, and (c) the perceived benefits and challenges of HR analytics. See Appendix 1 for the semi-structured guide used for the interviews. The interviews were conducted face-to-face or via Skype or phone. The interviews varied between 26 minutes to 1h 6 minutes in length. See Table 4. All the interviews were recorded and transcribed for later reference. All results are reported anonymously. The names of participating companies do not appear on this paper nor can the interviewees be directly connected to the case companies.

Table 3a. MNCs participating in the study

Company	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
Size	1,000-5,000	50,000+	1,000-5,000	500-1,000	50,000+	10,000-50,000	1,000-5,000	5,000-10,000
HQ location	UK	Finland	Finland	USA	Germany	Finland	Finland	Finland

Table 3b. Consulting firms participating in the study

Company	**I**	**J**	**K**	**L**
Size	<200	<200	<200	10,000+
HQ location	Switzerland	UK	USA	Germany

Table 4. Overview of interviews (continued on next page)

Interviewee	**Position**	**Tenure** (years)	**Experience of TM/analytics** (years)	**Duration of interview**	**Type**	**Language**
#1	Head of project delivery	10	20	44:41	Skype	English
#2	Talent acquisition specialist	0,33	15	46:46	Face-to-face (group)	Finnish

#3	HR analytics manager	4	3	46:46	Face-to-face (group)	Finnish
#4	Head of talent management	4	25	48:36	Face-to-face	Finnish
#5	SVP, Head of global HR	19	10	41:08	Face-to-face	Finnish
#6	Consultant	1	1,5	1:06:37	Face-to-face	Finnish
#7	HR director, global talent management	6	25	1:06:25	Skype	Finnish
#8	HR director, Americas	3	25	36:09	Skype	English
#9	SVP, Human resources	1,5	20	26:05	Skype	Finnish
#10	Head of leadership development	10	15	55:05	Face-to-face	Finnish
#11	Business people partner	7	5	49:11	Face-to-face	Finnish
#12	Chief strategy officer	1	18	44:36	Skype	English
#13	Principle project manager	2,5	5	50:03	Skype	Finnish
#14	Human capital leader	4	19	50:41	Phone	Finnish

The transcribed data was coded into keywords to gain a better understanding of the similarities and differences between the organisations. The findings were divided into the following 13 categories: talent definition; talent identification; talent identification challenges; performance, ability and affective component; tests used; tools used; data used; data challenges; HR ability to face analytics challenge; HR analytics benefits; HR analytics challenges; HR analytics future visions; and technology. The findings are presented in more detail in chapter 5, and discussed in chapter 6 with connection to the relevant theories presented in chapters 2 and 3.

4.4. Data quality

There are data quality issues that affect the reliability, generalisability and validity of a study (Saunders et al. 2012: 380). Reliability of a study refers to the extent to which a study can be reproduced. Studies with reliability have data collection techniques and analytic procedures that are able to produce consistent findings when repeated on another occasion or by another researcher. (Saunders et al. 2012: 192, Wilson 2014: 336.) Generalisability or external validity concerns the extent to which the findings of the study can be generalised to other relevant settings or groups (Saunders et al. 2012: 382). Internal validity refers to the extent to which a measure accurately reflects the concept it is supposed to measure (Wilson 2014: 337). In studies based on semi-structured and in-depth interviews, validity also refers to the extent to which the researcher has been able to gain access to the interviewee's knowledge and experience (Saunders et al. 2012: 382).

Since the data collection was done through semi-structured interviews, there is lack of standardisation which can lead to concerns regarding the reliability, generalisability and validity of the study (Saunders et a. 2012: 381). The interviews varied in length and some sub-themes were explored to large extent in some interviews while in others they were not discussed at all. This is in the nature of semi-structured interviews where the questions can vary from one interview to another based on the flow of the discussion (Saunders et al. 2012: 374–375). However, taking into consideration the purpose and nature of this study, it can be argued that it is not intended to be repeatable since semi-structured interviews reflect reality at the time of collection (Saunders et al. 2012: 382).

In attempt to increase the reliability of the findings, anonymity of reporting the findings was promised to the organisations and participating interviewees. This reduces the interviewee and participation bias which otherwise might lead to interviewees choosing not to reveal or discuss an aspect of the topic in fear of sharing sensitive information (Saunders et al. 2012: 381). This also affects the validity of the study as interviewees felt safe to share their ideas, knowledge and experience. Internal validity of the study was addressed by formulating and translating the questions to Finnish to make sure they were clearly understood. However, some terms lack well-established Finnish translations (e.g. talent) and much of the language used at MNCs regarding talent management is in English, which is why the participants were allowed to mix Finnish and English to express their thoughts and to describe different processes. Thus, the answers provided by the Finnish speaking interviewees are assumed to be as accurate as possible.

The cross-sectional data sample looks at the phenomenon at one short period of time. Moreover, the exploratory nature of this study aims to create an understanding of the phenomenon. As the field is rapidly developing, it is probable that repeating similar study at another time or with different participants would yield different findings. Furthermore, the findings represent the situation of a small sample of 8 MNCs and the views of 4 consulting firms. It is also important to note that most of the MNCs participating in the study are Finnish or European. While this multiple case study does provide more compelling evidence than a single case study, the findings of this study are not necessarily generalisable to larger populations or in other settings and should thus be interpreted with caution.

5. FINDINGS

This chapter presents the most significant findings from the empirical study. Since the case MNCs are at different stages of implementing HR analytics systems into their talent identification processes, there is a lot of variation between the case companies. First the findings related to talent identification process are presented and the findings on the data used in the process, followed by the findings on the perceived benefits and challenges of HR analytics and the expectations for future developments in the field.

5.1. Defining and identifying talent in MNCs

The MNCs participating in the study vary in size, sector and industry, which is why it is natural that their talent identification approaches and definitions differ. The eight MNCs reported five different approaches for identifying talent: by record (12,5%), by breadth (25%), by combination of record and breadth (37,5%), by level (12,5%) and by strategic position (12,5%). Half of the MNCs reported having an inclusive view on talent where every employee is seen as talent, while in reality their talent management appeared to be exclusive on those with potential for larger roles in the organisation. Despite the varying approaches to talent, the talent identification processes appear to be somewhat similar across the studied MNCs. Also, the skills, competences and personality traits used to describe talent were similar across the MNCs. The MNCs reported the following important attributes for talent: learning agility, cognitive ability, aspiration, motivation, flexibility, drive, leadership, emotional intelligence, and relationship building. However, there was variation on how much value is given to performance and the ability and the affective components. One MNC reported putting more emphasis on ability over performance, and two reported giving equal value to aspiration, ability and agility to learn. One MNC reported wanting to give more value to affective component and its behavioural aspects but in reality, performance and affective component are given the same weigh while another stated assigning equal importance to the dimensions. One MNC reported giving more importance to performance while two MNCs told they first look at performance but assign it equal value to the ability and affective components. Moreover, talent is seen as dynamic and an employee must constantly prove themselves to be talent.

> *"Once a talent is not always a talent – you must be able to prove each year why you deserve to be on this list."* – Company C

The talent identification process itself was reported to be very similar across the MNCs. Talent identification is a yearly process and done with connection to performance reviews. While guidance and support are provided by the HR function, the identification is done by direct supervisors who have responsibility over the process. After the first candidates have been identified, those people are evaluated more closely by HR and higher-level managers, typically board of directors or similar. This is done in the interest of calibration to have more consistent, transparent and objective talent identification. The identified talent is then placed in a talent pool for future development actions. One MNC however reported having a culture of internal recommendations which is something that did not come across anywhere else and suggesting the process to be less hierarchical than in the others. Moreover, they reported having multiple talent pools for different purposes and for talent at different stages. Other MNCs also reported segmenting their talent depending on the stage of talent.

> *"Speaking of talent pools, we want to have some in different stages as well, so that it works also in the long run. We don't want a situation where we have too many people who are ready to take on new and more demanding challenges, there needs to be a balance."* – Company B

There are a number of tools used to facilitate the talent identification process. The two most common tools mentioned by the MNCs were the 9-box grid and 360° evaluations. Only one MNC explicitly stated that they have abandoned the 9-box grid thinking. Another popular tool were surveys and tests, such as employee engagement survey, self-assessments, cognitive ability tests and personality tests. Such surveys and tests are bought as services from consulting firms. Two of the consulting companies taking part in the study provide such tools. The tools can be used by the MNC, but it is also possible that the consulting firm does the analytics for them. The survey and test results can be combined with the human resource information system (HRIS) data to gain insights from the workforce. 62,5% of the MNCs report using cognitive ability tests, personality tests and/or psychological evaluations during recruitment processes to support hiring decisions. However, the results are rarely used at a later time in talent management. Also, some reported using psychologists to observe their high potential groups in assessment centres.

Some of the challenges and concerns identified by MNCs, relating to talent identification, are the subjectivity and inconsistency of the process and those evaluating talent, and the impact of context, for example in a post-merger atmosphere. Some were also concerned with the narrow view of what talent might look like in their organisation and being stuck on faulty metrics. The consulting firms shared that common challenges their customers face include not having visibility into the talent data all in one place, and not understanding the impact of talent decisions, for example, on shifting demographic profile of the company. The consulting firms also pointed out that the organisations need to be very clear on their talent definition before applying analytics and technology into the talent identification process.

5.2. HR analytics and data in MNCs

All of the MNCs use HRIS software but not all of them apply it to their talent identification process. Also, the functionalities of the HRIS vary across the MNCs. One MNC told they use primarily spreadsheets to maintain talent information and others reported their process to be manual to a large extent even though data from the HRIS is used when identifying talent. Some also reported having to feed back the manually produced talent data into the HRIS. Some companies have HR BI systems purchased but the implementation is in progress and some of the modules of the systems are for the moment turned off. Many MNCs reported using Excel and PowerPoint to a great extent as a tool for processing the information despite having HRIS or HR BI software in the organisation.

The data used for the talent identification process consists mostly of HR master data and information gathered from performance reviews, talent reviews and surveys. Also, some MNCs reported using personal ability rating scores, financial metrics, key performance indicators (KPIs) and data from other information systems in the organisation. However, it is not common that combining information from different datasets is done by HR, rather it is outsourced to consulting firms or the IT function. Moreover, to support the results derived from data all of the MNCs highlight the importance of discussing the results, the role of emotional intelligence and the human side of the process.

> *"We have our HR IT team who can draw data from our HRIS and combine that information to business data and performance data, and that way combine data and create analytics and insights in ways that us mere mortals are unable to."* – Company F

5.2.1. Challenges related to data

There are several challenges that are related to data in talent identification. One of the most significant concerns the MNCs have is data privacy and the General Data Protection Regulation (GDPR), which is a regulation affecting the European Union from 25 May 2018 (EU GDPR Portal 2018). The MNCs want to be as transparent as possible with what they are doing with people data, as treating personal information is highly sensitive. Storing the data is not seen as a big challenge as large MNCs have the financial capabilities to take the necessary actions. Some report local regulations in subsidiaries to pose challenges. Also, what remains challenging according to the MNCs is getting data and ensuring its quality, accuracy and reliability, and maintaining it.

While the MNCs report that getting the data is one of their biggest challenges, the consulting firms suggest that the underlying issue might be that MNCs simply do not have good visibility on their data and struggle with making connections between the different processes and connecting what people say with what people actually do. Few of the MNCs identified the same challenge, as data processing, making connections between different data sources and data literacy were mentioned. Also, the static nature of the current data in use was identified as challenge. Another identified weakness was data usability as the MNC's HR BI system at the moment did not support combining different datasets from business functions to a desired extent.

> *"We're pulling together disparate datasets and one of the things we see in this space is people just don't have good visibility into their data, on average it lives in 10 different systems so even within HR or HR systems there are multiple systems or multiple ways to get data out of one system. So, we do a lot of bringing together multiple, disparate datasets and provide visibility into things like mobility or diversity or top talent. A lot of our customers and prospects say: 'oh my God, I have no visibility into my mobility data' and we're like 'you know you have it cause titles change and positions change' but there's not a good, easy way for them to see it."* – Company K

As a lot of the data used is based on test results and surveys, it is challenging to maintain up-to-date data. In fast-moving business environments broad competency studies, which some of the MNCs have done in the past, tend to be outdated by the time they are completed. However, the MNCs see competency analysis as an important part of talent management in expert organisations. The MNCs also struggle with getting data on some of the most significant dimensions of their talent definition, such as learning agility. The MNCs are concerned whether they collect the right data, whether they have the right metrics and whether they are asking the right questions. The consulting firms point out that using survey data mixes facts with feelings, perceptions and what people believe to be the right answer. However, the use of big data in HR analytics seemed to be intimidating to MNCs. When asked, some of the MNCs recognised the potential of using internal social media channels such as Yammer as a source for talent data while others did not see the value of such network analysis. Most of the consulting firms would call for such network analysis and further analysis based on different types of tracking. The MNCs were not enthusiastic about tracking their workforce for talent purposes as they feared it might be too sensitive of an issue for the employees from privacy perspective.

> *"I think what the people management departments in many areas have not understood is that the truly relevant – I mean, of course it's interesting to ask people what they think – but the truly interesting data doesn't come from asking for the perception of people, but actually looking at what do people really do."* – Company I

5.2.2. Perceived benefits

All MNCs view using HR analytics as a growingly important part of HR and talent management. All participants view the biggest benefit of applying HR analytics to talent identification and management to be fact-based decision making and argumentation. Other important benefits of HR analytics are objectivity, transparency and decreased ambiguity which are gained from the fact-based decision-making and will allow organisations to address issues that might relate to diversity. Some MNCs reported having diversity initiatives to address age, gender and nationality in their talent pools and saw the benefits of HR analytics related to their efforts.

> *"I think it is increasingly important that we evaluate our talent pool critically. We have certain development areas – we don't have enough women in the organisation, especially the higher we get on the organisational ladder. Our organisation remains too Finnish-based – we need more international talent. And we as HR must take this as our agenda and be the spotlight in the organisation to address these issues and keep these in mind when new positions open up and so on. [---] We're identifying junior talent, but another important category is senior talent – if we're meant to retire when we're 80 then we simply cannot stop identifying talent at 35."* – Company F

The potential of predicting analytics is widely recognised among the MNCs. The MNCs believe that HR analytics can help to gain visibility into the future by predicting movement and activities within the workforce, and to predict success, which will enable the organisation to proactively manage careers to increase efficiency. The MNCs believe that HR analytics can help them to build the right people to become the future leaders. Moreover, by combining data from different systems such as HRIS, ERP and CRM, organisations will be able to understand correlations and to allocate their resources better. Some of the MNCs perceive that using HR analytics will eliminate the risk of error caused by manual work, while others believe that analytics would not affect the accuracy of identification. Furthermore, consulting firms list benefits such as being able to focus on retaining the right people, having more scientific approach, not repeating the same mistake year after year, not only identifying those performing on a high level right now but with the potential to grow for bigger role and higher performance in the future, and gaining visibility into competency gaps in the organisation.

> *"It would extremely valuable to us if we could say to person X 'hey, do you need new career challenges to be more engaged and productive here' – we could proactively manage careers and give our employees the right challenges when they need and are ready for them."* – Company A

> *"HR analytics would reduce the amount of manual work which would eliminate the risk of error caused by that."* – Company H

5.2.3. Perceived challenges

Many of the challenges the MNCs reported relating to HR analytics had something to do with data. The MNCs were concerned how can they get the tacit knowledge embedded in people into the system. Another category that got attention was mistrust in technology as the MNCs identified a learning curve needed for fully adopting HR analytics into their organisations and some pointed out that managers will be able to justify around the data when they want to. Many were also concerned that if the processes are automated and the output data is blindly accepted, the organisations are risking losing compassion and valuable information gained through discussions. The MNCs do not see that data and HR analytics alone will give 100% accurate insights as there are several human factors behind people and AI cannot substitute for the human factor in interactions, accentuating the importance of emotional intelligence. One MNC stated that talent cannot be based on analytics alone and that more complete evaluation and feedback is needed to gain as objective results as possible by combining information from various sources.

> *"Analytics alone cannot be the only factor determining this person's ability, there has to be more comprehensive evaluation and feedback so that it doesn't happen so that it is based on one person's, supervisor or colleague for example, evaluation which can be overly positive or negative. We must use multiple tools and combine the information to gain an objective overview of the situation."* – Company C

One MNC highlighted the meaning of context in skills and competences which the computer software is unable to take into the equation. One of the consulting firms was along the same lines as they pointed out that the same models do not work for every company as a lot is dependent on industry and company type. The MNCs were concerned with trusting the algorithm but they also doubted their own ability to use this type of technology. One MNC was concerned that not knowing what you really want and need would lead to accidentally using the wrong parameters to screen candidates would rule out the individuals that the company is actually interested in. One MNC stated that talent is a complex and developing concept, making it difficult to apply analytics to it. Moreover, one of the consulting firms pointed out the difficulty of combining natural science with social science and taking consistency or inconsistency into account when automating.

> *"[--] part of the challenge is you're dealing with social science, not science. You can't just ingest all my data and say 'hey, let me predict stuff about your organisation!' It doesn't quite work like that, so I think on the social science side IBM Watson faces a lot of the same challenges as everybody else and it requires human intervention to understand the data and what the machine is trying to tell you and to see if it's picking up the right ques."* – Company K

5.2.4. HR ability to take on the data challenge

All MNCs recognised that HR needs different kind of capabilities in order to take on the HR analytics challenge. The MNCs had identified what the consulting firms believe to be the biggest challenges: HR is not used to working with data and such fast moving environment, and HR does not have the analytical capability to understand what conclusions can be drawn from which data, to conduct analysis and to interpret data. Taking on the HR analytics challenge takes a different type thinking and a different type of person that's in typical HR. There is need for interdisciplinary competencies from domains beyond HR, such as data scientists, mathematicians, statisticians, and psychologists. Also, there is need for more deeper understanding of different business functions in the organisation, which can be solved by building the HR analytics team of experts from different business functions and areas under a global centralised HR function.

> *"I would say we could have the capability if we take the time. I'm not sure everyone in HR has the capability, but we can have that capability, it's a different type of thinking and a different type of person that's typical in HR."* – Company D

> *"HR analytics is clearly its own field of expertise and HR needs to understand its possibilities. It becomes increasingly important to understand the big picture so that you can articulate what you want and what you think you could find in the data."* – Company G

5.2.5. Visions for the future

One MNC expressed their dissatisfaction with the current commercial HR BI solutions and envisioned greater things for the future. Their vision is to implement standardised dashboards and visibility on urgency of chain, and to bring IBM Watson more into their people analytics. They viewed combining information from different sources as a key to achieving better analytics. Other MNCs acknowledged the growing of the field and the need for it and saw AI as a future trend. The MNCs view more advanced analytics as a gateway to predict talent and to build stronger organisations through excellent workforce and retaining the right people and letting go of those who are not performing at their best. Multiple channel strategies are seen as the answer to find the right talent and focus groups and to identify characteristics that predict high performance. Some MNCs identify the potential in leveraging the Internet of Things (IoT) to gain insights on for example employee wellbeing to make proactive talent management decisions.

> *"As we know in medicine, soon we're getting chips that will analyse our blood sugar and how you are doing, and should you get some rest and so on – this I believe will happen faster than any of us can predict. So, if there is a link, even though its physical, could we draw some insights from it to my job performance as a marketing manager or something, such as wellbeing and stress levels."* – Company E

One MNC stated that analytics on behavioural data is more accurate than self-provided and "branded" information put online, making it more reliable to use in talent identification. This thought was supported by the consulting firms as according to one, the truly relevant data comes from looking at what people do. The consulting firms were also keen on doing social network mapping and relationship network mapping to gain insight on talent. Moreover, the consulting firms see great potential for leveraging IoT and using unstructured data to better understand workforce, and big data, data mining and predictive analytics are seen as the driving forces of the future.

> *"Everyone is talking about big data, and data mining and predictive data are currently hot topics, so I think that it creates sort of positive pull for the whole field so that companies actually want to invest time and money into getting these things right."* – Company L

6. DISCUSSION

This chapter takes the findings of the empirical study and connects them to the literature review in chapters 2 and 3. First, the concepts of talent and talent identification are discussed, followed by discussion on data analytics and HR.

6.1. Defining and identifying talent in MNCs

Research shows that talent identification in MNCs is a structured and hierarchical process (Pepermans et al. 2003, Dries & Pepermans 2008, Nijs et al. 2014). The findings of the study support this, as it is described to be a yearly process which is done in connection with performance reviews. Even though direct supervisors have responsibility over the process, it is common across the MNCs that HR provides a framework for the managers to follow, and support in carrying out the process. Only one of the case companies shows signs of a less hierarchical process as they reported having a culture of internal recommendations. After the first pool of candidates is obtained, HR takes more control over the process and calibrates the output of the talent review with higher level managers. Once talent has been identified, it is put in talent pools, which is a group of employees who have been identified as talent (Tansley 2011). Dries and Pepermans (2008) posit that it is common for MNCs to have talent pools for different purposes. While senior, junior, female and international talent were discussed in some MNCs, only one MNC stated clearly having talent pools for different stages of talent. Merely segmenting by age does not count as segmenting talent by its stages. For instance, someone might switch careers at age 40 and have similar job-related competencies as someone in their late 20s, making their stage the same even though there is a large age gap. Also, as was accurately pointed out by one interviewee, careers are expected to last longer and longer as retirement age is pushed back, which means that identifying talent should go well beyond 35-year-olds.

Silzer and Church (2010) maintain that talent identification should start by defining talent and potential, and deciding on what categories to consider. The talent multiplication model presented by Cheese et al. (2008) is a great framework for MNCs to base their talent management on. However, talent is a complex concept which is recognised by the MNCs participating in the study as well as by academia. There are a number of different approaches and perspectives to talent and how to define and identify it. The MNCs participating in the study all see that identifying talent is important for their organisation.

However, when asked how they define talent, the answers varied from clear definitions to vague descriptions. Collings and Mellahi (2009) suggest that organisations struggle with talent management due to lack of consistent definition and clear conceptual boundaries in the field. This can be argued to be true to some extent as the vaguer the definition an MNC had, the more concerned they seemed to be on whether they are making the right decisions, using the right metrics or collecting the right data. The meaning of context arose more than once during the interviews. It is clear that for companies in different industries and sectors the need for talent is different, yet the definition of talent remains somewhat similar. This could suggest that there are universal characteristics of talent or that companies do not have clear enough understanding of what talent in their organisation looks like and end up with vague definitions. Some of the MNCs express their concern on making the wrong talent decisions – identifying someone who turns out not to be talent or failing to recognise someone as a star employee. The former situation can lead to wasted resources and the latter to missed opportunities.

Collings and Mellahi (2009) further suggest that talent management activities should be rooted in the organisation's strategy. Such link between talent definition and company strategy was missing in most of the definitions provided by the MNCs. However, one MNC has their talent definition clearly built around the company strategy and the starting point for their talent identification is through identifying the strategic key positions in the organisation. Silzer and Church (2010) identify approaches to defining talent, one of which is by strategic position. While one MNC has their talent definition embedded in the company strategy, most of the MNCs in the study have chosen to focus on the "A players" which many scholars have researched (eg. Tansley 2011, Lepak & Snell 2002, Nijs et al. 2014).

According to Tansley (2011) talent is often associated with individuals who have the potential to reach high levels of performance in the organisation and who show the most potential for moving up to more senior roles. The MNCs in the study have these incorporated in their talent definition to some extent. Further, Silzer and Church's (2010) approaches to identifying talent include by role and by level, which relate directly to Tansley's (2011) findings on what is associated as talent. One of the MNCs has talent by level incorporated in their talent definition, which means that to them a high potential individual is someone who can move up one or two organisational levels.

Another way to anchor talent definition is by record (Silzer & Church 2010) which is based on past performance and it is among the most popular ways of defining talent among the MNCs in this study. Half of the MNCs include consistent and high performance as one of the most important defining traits of talent. However, Nijs et al. (2014) maintain that performance should not alone be used as defining characteristic for talent and promote the use of affective and ability components. Of the MNCs that value performance highly when identifying talent, 75% also attach defining talent by breadth to their definition which refers to capacity to take on larger and more complex tasks (Silzer & Church 2010). The MNCs that assign both breadth and record to their talent definition are closer to the talent definitions agreed on by researchers where not only performance is valued but also ability to make a positive difference in the organisational performance (Chartered Institute of Personnel and Development 2007) and the human capital of the individual which rates high on strategic value (Lepak & Snell 2002). 25% of the MNCs in the study assign the capacity to take on larger and more complex tasks as their primary way of defining talent aligning with the findings of Nijs et al. (2014) – ability and affective components should be emphasized over performance results.

Some of the MNCs state that they rate performance and ability and affective components equally important when they are identifying talent. For some of the MNCs, high performance is a precondition before someone can be taken into consideration to be labelled as talent. As Nijs et al. (2014) posit, the purpose of talent identification is not only to recognise talent already expressed within the organisation but also to identify those who have potential to excel in different roles in the future. This means that making high performance into a precondition for talent might be a mistake as someone with great potential might be misplaced in the organisation and unable to reach their full performance potential. Potential, which is associated with talent, implies there is possibility for an individual of becoming something more than what they currently are (Silzer & Church 2009b). However, recognising the potential in someone who is not a consistent high performer is likely to be very challenging. Moreover, understanding the underlying traits, characteristics and behaviours of potential – or talent – is extremely difficult, yet it is necessary before beginning to apply analytics into the equation.

Another point of discussion is whether talent is innate or if it can be developed later in life. Innate talent is built around ability and personality, which are stable competences. Talent that can be developed later in life is built around skills and knowledge. (Leiba-O'Sullivan 1999.) Many of the MNCs use tests to some extent to determine the cognitive ability and personality of individuals which suggests that stable competences are seen as an important dimension of talent. Also, the consulting firms offer such tests which further suggests their view of talent is also based on stable competencies. While the importance of stable competences is recognised, it is not uncommon for the MNCs to value dynamic competences over stable ones when making decisions under time pressure. Such ad hoc approach has also been identified by McDonnel et al. (2010) who found that MNCs are often inconsistent with their talent management. However, it is also common to combine stable and dynamic competences when identifying talent, as the two types of competences are closely intertwined and support each other (Leiba-O'Sullivan 1999). This is very common in the case companies as many of the MNCs identify as one of the most important competency was learning agility. Literature recognises learning ability as a competence relating to the cognitive meta-competence (Tansley 2011, Brown & McCartney 1995, Le Deist & Winterton 2005). The way learning agility is described by MNCs suggests it is close to interchangeable with learning ability. The MNCs see that individuals with learning ability have potential to grow their dynamic competences.

Some of the MNCs have adopted an approach where they posit that everyone is talent, but some have the potential for more which will determine whether they are considered for being part of talent development programs and talent pools. The MNCs might consider they have an inclusive talent management approach while in reality their approach is exclusive. These companies value the affective and ability components when identifying high potential from their workforce. The MNCs look for the cognitive intelligence, emotional intelligence and social intelligence competencies in these individuals as a way of determining whether they have potential. Also, some of the MNCs use some type of psychometric testing to support their talent decisions. Boyatzis (2008) suggests that competencies are the behavioural manifestation of emotional, social and cognitive intelligence, which is why using psychology professionals to evaluate the behavioural aspects is well justified. Moreover, psychological testing is well suited for determining personality of the candidates as personality – a stable competence – relates to the affective component of talent. According to Nijs et al. (2014) the affective component consists of non-intellectual attributes such as personality, interests and motives, which affect individual performance. Many of the MNCs recognise the importance of these as determining factors for potential. One of the most important dimensions of personality is agreeableness, as most MNCs highlight the behavioural side of performance – how results are achieved. Acting according to company values and not putting others down is valued, as are also personality traits associated with agreeableness: flexibility, good-natured and cooperative (Barrick & Mount 1991).

The affective component also includes the measurement of potential, as personality alone does not predict talent (Nijs et al. 2014, Tyler 2011, Silzer & Church 2009a). The ways the MNCs in this study measure and detect potential in their workforce varied a lot. Some use cognitive ability tests, personality tests and psychological tests to draw conclusions on the potentiality of an individual while others base their measurements on sets of questions that supervisors answer based on their perception of the employee. The performance-potential –matrix is used in most of the MNCs as a tool to measure potential. The tool has been criticised as not being adept for measuring potential (Lumme-Tuomala 2017) and one of the MNCs has completely abandoned the 9-box thinking. The criticism of the tool is justified to some extent as companies use the tool for measuring potential and for that it is not adequate. However, the tool can be highly useful when categorising employees for example into different talent pools. Instead of relying on the performance-potential –matrix or the potential pyramid by Robinson et al. (2009), companies should find ways to better incorporate the measurement of potential. Using sets of questions to determine potential in an employee makes the process highly subjective. Lumme-Tuomala (2017) suggests that potential can be determined by person’s ability, engagement and aspiration to rise and succeed in the organisation. This is why the combination of different tests is likely to be more accurate determinant of potential.

According to Fernández-Aráoz (2014) potential can be predicted from curiosity, insight, engagement and determination. Some of the consulting firms participating in the study provide employee engagement surveys. Such surveys hold a great amount of value for the MNCs and could be applicable when identifying talent. However, MNCs do not necessarily make the connection with survey data and talent data. Also, the reliability of such tests can be questioned as well, since some consulting firms preferred looking at behavioural data over survey data, while others argued that the survey results are sufficiently accurate.

6.2. Data, analytics and HR

While organisations create a lot of data during their normal course of business, there is much data that optional to create as it is non-essential for the business (Pape 2016). Some MNCs reported that they do not see the benefit or the return of collecting certain types of data on their employees. At the same time, HR functions are asking whether they are collecting the right data when the question should be whether they are looking at the right places. It is suggested that the MNCs cannot find the data they are looking for since HR data alone can live on average in 10 different systems. It is likely that there is enough data created in the organisation for talent identification needs but the organisational structures are preventing its efficient use across functions. The data that the MNCs currently use for talent identification is very static by nature. Large MNCs have a lot of data, not only on their employees but other business data as well, which is why the MNCs in the study do not see storing the data as such big of an issue, contrary to what is suggested by Sivarajah et al. (2017).

6.2.1. The challenges

Cascio and Boudreau's (2008) LAMP model for evidence-based talent identification is a simplified framework to guide MNCs to implement analytics into their organisations. The components of the model are the right analytics, the right measures, the right process, and the right logic. Taking the model to MNCs would require a lot of work, as majority of the case companies were concerned about whether they are measuring the right things or have the right data, how to leverage value from their data and if they are even looking at the right things. Similar remarks were made by the consulting firms as they questioned HR's ability to process data and make connections between different datasets to gain insights.

The concept of big data appears to be too large and complex for the case companies and especially for their HR functions to manage. Even the concept of data seems to be too difficult for HR to handle. The MNCs experience similar challenges related to data as are identified in research to be associated with big data. For example, Sivarajah et al. (2017) identify big data management challenges such as privacy, security, data governance, information sharing, data ownership and ethical issues of data usage. Most MNCs in the study reported having these issues with their HR master data, which compared to big data is simple and well-structured. The MNCs are unable to see how they could in practice leverage insight from multiple sources such as company's internal social media channels or tracking combined with the data they are currently using. According to one of the consulting firms, MNCs are too afraid to do anything with data that they think might upset employees, which is why they fail to gain insights on the truly relevant data. This seems to be true to some extent in the case MNCs as privacy issues are associated with the static HR master data. The consulting firms also suggest anonymising data for employee protection. Moreover, one consulting firm believes that the person behind the data is irrelevant as the big picture and trends shared by certain employee segments are more interesting and useful for the organisation.

It can be argued for and against why HR analytics should or should not be performed in the HR function. One of the consulting firms is a strong advocate for maintaining anything related to HR and people within the HR function while others were more open to centralised data functions. Locating the HR analytics function into a centre of expertise where the rest of the BI functions are would make sense, as the MNCs and the consulting firms all agree that HR is not capable to process large amounts of data. Minbaeva (2017) and King (2016) are along the same lines as they identify a gap in HR's capability and question the ability of HR function to produce analytics and insights for organisational gain. MNCs identify that it takes a different type of thinking and a different type of person than is typical in current HR to take on the data challenge. A combination of different kinds of interdisciplinary capabilities is likely to be the answer, such as data scientists and engineers, psychologists, statisticians, mathematicians without forgetting the HR professionals. Taking on the data challenge requires a data-oriented mindset with understanding of behavioural aspects and emotional intelligence.

It is not only that HR is not ready for the big data challenge, but neither are the HR BI systems that are commercially sold at the moment. The current systems are built on relational databases which require the user to know which data to input and what to ask (Pape 2016). This is problematic as HR commonly lacks the capability to understand what datasets to combine to gain valuable insights for the organisation. Some of the MNCs claim that the issue is getting the data, while van Knippenberg et al. (2015) argue that it is not about finding and cumulating data but about having long enough attention span to make sense of the data. Efficient methods for analysing data are extremely important and while artificial intelligence and machine learning are ideally suited for processing and extracting information from large, connected datasets and big data (Kitchin 2014:100-101), this has not yet reached the commercial solutions of HR BI. The issue is not that the HR BI software does not have efficient enough computation but rather the premises attached to the use of the software. It is clear that HR struggles with asking the right questions and making connections with different datasets, which is why such software does not support HR in their work. HR has to know what to ask from which datasets before they can use their expensive HR BI systems and since they lack the analytical capability, the software is not used efficiently.

Most of the MNCs use HR master data and information they gather from performance reviews, talent reviews and surveys when they are identifying talent. These data are stored in the same HRIS which makes them easy for HR to access. However, the created information does not necessarily produce valuable insights. Angrave et al. (2016) and Pape (2016) argue that more valuable insights can be extracted when big data systems are leveraged instead of relational databases. Moving beyond the HR master data and review and survey data is not on the plate at the moment for most of the MNCs. Many MNCs mentioned data ownership as a challenge and used the ownership of HR master data as an argument to maintain HR analytics within the HR function. As Angrave et al. (2016) suggest, traditional silo structures affect the data management mentality, preventing HR related data from being combined with data from other organisational functions. This results in loss of valuable insight on the data. Moreover, according to Pape (2016), it is important to understand what type of data are being created and what data are needed to answer the analytics problem at hand. Some of the MNCs expressed that they are not sure of what the data in their HRIS consists of and whether there is some business data behind the talent data. Since HR does not know of which components their talent data are constructed, it becomes increasingly difficult to ask the right questions.

According to Minbaeva (2017), MNCs do not know how to begin the in-house development for bridging the capability gap in HR. One of the consulting firms suggests that MNCs should abandon the idea that everything has to be built in-house. Rather, companies should seek out the help of HR analytics professionals when beginning to implement people analytics. However, even if an MNC has more capability than the average to ask the right questions, the challenges of the HR BI systems and data usability remain. One of the case companies with more advanced systems for HR analytics stated that their HR BI system does not support combining datasets from different business functions to a desired extent. This would suggest that the silo structures are not only affecting the thinking of the HR function, but also the design of the commercial software. The MNC in question has vision to combine business data with their workforce planning data to produce predictive analytics that provide insights on the next two years of what type of competencies are needed in different business areas and regions.

6.2.2. The benefits and future applications

Research shows that predictive analytics hold great potential for organisational management and development (Fitz-Enz 2014: 2) and this has not gone unnoticed by MNCs. However, based on the empirical evidence from the study, it appears that currently most MNCs are capable of descriptive analytics at their best. Most MNCs are looking at HR analytics as a way of supporting their argumentation and decision-making with facts and statistical evidence. The benefits of gained objectivity, transparency and decreased ambiguity in talent decisions, which are mentioned by many of the case companies, suggest that current organisational practices are defective. A lot of the decision-making is based on perceptions, opinions and unconscious biases even though calibration is an important part of the talent identification process in the MNCs. Using analytics can revel organisations mistakes they have repeated over the years in their talent decisions and enables organisations to address these pressure points.

Some of the MNCs envision using HR analytics and different types of talent pools to help the organisation respond to talent needs more precisely and efficiently. For example, if a critical person leaves the organisation, the MNC is able to quickly respond to the change with the right measures. However, this can be achieved also by having comprehensive understanding of the competencies and skills of the workforce which does necessarily require analytics. Being able to identify critical points in the organisation, that is, who can we not afford to lose and are they at risk leaving, is more valuable for MNCs. This could be obtained with predictive analytics, yet only few of the MNCs are currently working with what could be called predictive analytics.

Predictive analytics have the potential to create insights about the future through forecasts and statistical modelling (Kitchin 2014: 101, Sivarajah et al. 2017). While the most efficient methods for predictive analytics include applications of artificial intelligence and machine learning, (Kitchin 2014: 100–101) it is possible to leverage predictive analytics without such technologies (Sivarajah et al. 2017). None of the MNCs had artificial intelligence helping their internal talent identification, but many see it as a future trend. Also, as was pointed out by one of the consulting firms, combining natural science with social science can be problematic. Combining social science with natural science requires human intervention to make sure the machine is picking up the right ques, which is why unsupervised machine learning is not applicable to talent identification. Previous research has shown that most value can be created through supervised machine learning techniques, more specifically different data mining techniques such as decision trees, artificial neural networks and support vector machines (Jantan et al. 2009c; 2010).

It is commonly recognised among academics that domain expertise is needed when working with machine learning algorithms to justify the meaning of extracted knowledge (eg. Chien & Chen 2008, Wei & Xiaolin 2011, Bell 2014). Even though HR has a lot of domain expertise in talent, the sources of data that HR currently sees fit for talent identification suggest it is likely that they still lack some of the analytical capabilities needed to make sense of the output data. HR seems to be stuck in the idea that talent identification has to be based on the talent review process. Consulting firms maintain that looking at behavioural data over survey data yields more valid results for talent identification. Only one MNC recognises the value of using behavioural data over branded information people knowingly create online. Chien and Chen (2008) also propose using behavioural data and connecting that to personnel profiles by applying a decision tree algorithm. Jantan et al. (2009b) found that decision tree algorithms can reach up to 95.14% accuracy in forecasting talent.

Some MNCs in the study see that manual processes or their current practices produce accurate enough results and one even stated that they believe they can get close to 100% accuracy. However, research has shown that there is high uncertainty associated with manual talent identification processes which can be overcome with automated systems (Goonawardene et al. 2010). Artificial neural networks can reach on average the accuracy of over 91% (Zhang et al. 2013, Wei & Xiaolin 2011, Goonawardene et al. 2010). Based on the case companies' HRs' habit of using HR master data, performance reviews and various survey and test scores to support their talent identification, it seems that adapting artificial neural networks could be the answer to automating the process. Huang et al. (2004) propose identifying managerial talent with artificial neural networks based on individual traits, such as capability, motivation and personality, and managerial skills, such as conceptual skills, interpersonal skills and technical skills. The model is capable of learning very complicated relationships (Huang et al. 2004) and avoiding human error by being able to self-organise, self-adapt and make decisions in paradoxical knowledge environments (Wei & Xiaolin 2011). Thus, applying an artificial neural network model to the data MNCs currently use could provide more accurate outputs.

7. CONCLUSIONS

The purpose of this thesis was to examine to what extent organisations create, capture and leverage value for their talent identification process through HR analytics. Three research objectives guided the study, i) How is talent defined in MNCs?, ii)What data analytics tools are applicable for talent identification?, and iii) How are analytics used in MNCs to benefit the talent identification process? The empirical evidence for the study was gathered through interviews with 14 professionals knowledgeable about HR analytics in 8 MNCs and 4 consulting firms.

The findings of the study were aligned with those in previous research in terms of how talent is defined and identified in MNCs. The definitions of talent varied between the MNCs and were not always very clear. MNCs experience uncertainty when identifying talent which is likely to result from unclear definitions. Applying analytics in organisations that are unable to articulate what they consider as talent is probably a waste of time since analytics are based on statistical modelling. It is very challenging to build models without understanding what the model is trying to achieve and what factors characterise the desired output – in this case *talent*. Once the stable determinants of potential are identified, defining talent by using stable competences can be an asset when programming talent software as the parameters are unlikely to change. However, the ways employees manifest stable competences in their work may vary depending on the type of organisation and the type of work the employee does, which is why it is important to take context into consideration when building such models.

Previous research has identified a number of big data tools with potential to address personnel selection problems through classification. However, it does seem that their use is still the future for the MNCs that have just started, or plan to start engaging with people analytics. The majority of MNCs are not currently using very sophisticated tools in their talent identification process. The talent identification processes remain for the most part very hierarchical. With big data tools the talent identification process could be made into a continuous and more agile process that improves the overall flow of the workforce. Moreover, the use of big data tools could increase the accuracy of talent decisions which would address the concerns on making the wrong talent decisions.

With more diverse use of data and combining data from different sources, MNCs could be able to identify not only those who perform at a high level but also those who have the potential to do so in the future. With the current methods, MNCs need new ways to measure the affective and ability components of talent and not be fixed on the performance aspect. More importantly, MNCs need to understand how the characteristics of talent for their organisation are manifested in employees' behaviour and how to interpret behaviour into the different dimensions of talent.

The current HR BI systems are challenging due to the assumption that HR knows what questions to ask and from which datasets. Scholars agree with what was shown in this study – the HR function does not have the analytical capabilities and technical expertise to combine data from disparate datasets and draw insights. However, the lacking analytical capabilities of HR are not the only challenge since data privacy continues to challenge organisations and can further raise ethical questions related to data usage. The changes in EU regulations is making companies to hit the brakes with some data related initiatives as organisations are unable to predict the effects of the upcoming regulation. In addition to data privacy, data maintenance is challenging to HR as much of the data is based on reviews and observations from managers, instead of being created automatically.

7.1. Managerial implications

MNCs should focus on understanding the underlying attributes of talent in their organisation. Companies should put effort into being thorough when implementing HR analytics into the organisation because otherwise the results will not create the desired value. Replacing old concepts with new buzzwords does not necessarily change what the organisations aim to measure. Instead of reinventing the wheel and rebranding the HR function, MNCs should make sure they actually achieve what they set out to do and get the return they want from their investments.

MNCs should focus on building a team with diverse backgrounds to do people analytics. HR alone clearly lacks the capability to face the data challenge, but data and software engineers cannot do it alone either. In addition to HR people and data scientists, there needs to be a link that can facilitate the communication. Such link would ideally have interdisciplinary background in data science and HR and organisational development. Experts such as statisticians are also a great asset to the team since they have the expertise in statistical modelling. In addition to these, the team needs to have an understanding of the business in order to be able to address its needs and add value to the whole organisation.

7.2. Limitations and suggestions for future research

There are three main limitations in the study. First, the case organisations represent a Finnish point of view, excluding other parts of the world. Second, the distribution of participants between HR professionals and analytics or data professionals was imbalanced as over half of the interviewees were HR professionals. Third, the size, industry or sector of the MNCs were not controlled. Also, the focus was on identifying talent within organisations leaving talent acquisition out of scope as the idea was to understand they ways companies can leverage their existing workforce with the existing data in the organisation. Big data tools for personnel selection are also used in the field of talent acquisition. There are not many studies on using big data tools specifically in talent identification within organisations

In the future studies, comparison between MNCs of approximately the same size from different geographical locations and more in-depth data collection could provide more accurate image of the state of HR analytics in MNCs. Including a balanced mix of different type of experts could enable more valuable insights. The applications of big data tools for internal talent identification could be further explored.

LIST OF REFERENCES

Adriaans, Peter & Dolf Zantinge (1996). *Data mining*. Harlow: Addison-Wesley. 158p. ISBN 0201403803.

Angrave, David, Andy Charlwood, Ian Kirkpatrick, Mark Lawrence & Mark Stuart (2016). HR and analytics: why HR is set to fail the big data challenge. *Human Resource Management Journal* 26:1, 1–11.

Aral, Sinan, Erik Brynjolfsson & Lynn Wu (2012). Three-Way Complementarities: Performance Pay, Human Resource Analytics, and Information Technology. *Management Science* 58:5, 913–931.

Barrick, Murray R. & Michael K. Mount (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Psychology* 44:1, 1–26.

Beechler, Schon & Ian C. Woodward (2009). The global "war for talent". *Journal of International Management* 15, 273–285.

Becker, Brian E., Mark A. Huselid & Dave Ulrich (2001). *The HR scorecard: Linking people, strategy and performance*. Boston: Harvard Business School Press. 235p. ISBN 978-1-57851-136-5.

Bell, Jason (2014). *Machine learning: Hands-on for developers and technical professionals*. New York: Wiley. 380 p. ISBN 1-118-88949-5.

Boston Consulting Group (BCG) (2008). Creating People Advantage: How to Address HR Challenges Worldwide Through 2015.

Boudreau, John W. & Peter M. Ramstad (2005). Talentship, Talent Segmentation, and Sustainability: A New HR Decision Science Paradigm for a New Strategy Definition. *Human Resource Management* 44:2, 129–136.

Boudreau, John W. & Peter M. Ramstad (2006). Talentship and HR Measurement and Analysis: From ROI to Strategic Organizational Change. *People and Strategy* 29:1, 25–33.

Boyatzis, Richard E. (1982). *The competent manager: A model for effective performance*. New York: Wiley. 308p. ISBN 0-471-09031-X.

Boyatzis, Richard E. (2008). Competencies in the 21st century. *Journal of Management Development* 27:1, 5–12.

Briscoe, Jon P. & Douglas T. Hall (1999). An alternative approach and new guidelines for practice. *Organizational Dynamics* 28:2, 37–52.

Brown, Reva Berman & Sean McCartney (1995). Competence is not enough: meta-competence and accounting education. *Accounting Education* 4:1, 43–53.

Buckingham, Marcus & Richard M. Vosburgh (2001). The 21st Century Human Resource Function: It's the Talent, Stupid! *HR. Human Resource Planning* 24:4, 17–23.

Cappelli, Peter (2008). Managing Uncertainty: Talent Management for Twenty-First Century. *Harvard Business Review* 86:3.

Cascio, Wayne F. & John W. Boudreau (2008). *Investing in people: Financial impact of human resource initiatives*. Upper Saddle River, N.J.: FT Press. 324 p. ISBN 978–0–13–239411–6.

Chartered Institute of Personnel and Development (2007). *Talent management*. Available online: <URL:http://www.cipd.co.uk/NR/rdonlyres/B513502C-8F42-419C-818C-D3C12D87E0D3/0/talentmanage.pdf>.

Cheese, Peter, Robert J. Thomas & Elizabeth Craig (2008). *The Talent Powered Organization: Strategies for Globalization, Talent Management and High Performance*. London and Philadelphia: Kogan Page. 276 p. ISBN 978-0-7494-4990-2.

Chien, Chen-Fu & Li-Fei Chen (2008). Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. *Expert Systems with Applications* 34:1, 280–290.

Collings, D.G. & K. Mellahi (2009). Strategic Talent Management: A review and research agenda. *Human Resource Management Review* 19:4, 304–313.

Dries, Nicky & Roland Pepermans (2008). "Real" high-potential careers: An empirical study into the perspectives of organisations and high potentials. *Personnel Review* 37:1, 85–108.

Dries, Nicky & Roland Pepermans (2012). How to Identify Leadership Potential: Development and testing of a consensus model. *Human Resource Management* 51:3, 361–385.

Dursun, Mehtap & E. Ertugrul Karsak (2010). A fuzzy MCDM approach for personnel selection. *Expert Systems with Applications* 37:6, 4324–4330.

Eriksson, Päivi & Anne Kovalainen (2016). *Qualitative Methods in Business Research*. 2nd ed. London: SAGE. 363 p. ISBN 978–1–4462–7339–5.

Ertel, Wolfgang (2011). *Introduction to artificial intelligence*. New York: Springer. 316p. ISBN 978-0-85729-298-8.

EU GDPR Portal (2018). GDPR Portal: Site overview. Available from Internet: <URL:https://www.eugdpr.org/>.

Fan, Chin-Yuan, Pei-Shu Fan, Te-Yi Chan & Shu-Hao Chang (2012). Using hybrid data mining and machine learning clustering analysis to predict the turnover rate for technology professionals. *Expert Systems with Applications* 39:10, 8844–8851.

Fernández-Aráoz, Claudio (2014). 21st-century Talent Spotting. *Harvard Business Review* 92:6, 46–54.

Fitz-Enz, Jac (2010). *The New HR Analytics – Predicting the Economic Value of Your Company's Human Capital Investments*. New York: AMACOM in association with GSE Research. 369 p. ISBN–10: 0–8144–1643–8.

Fitz-enz, Jac & John Mattox (2014). *Predictive Analytics for Human Resources*. Wiley. 149 p. ISBN 1-118-89367-0.

Goonawardene, Nadee, Shashikala Subashini, Nilupa Boralessa & Lalith Premaratne (2010). A Neural Network Based Model for Project Risk and Talent Management. In: *International Symposium on Neural Networks. Springer, Berlin, Heidelberg,* 532–539.

Harris, Stacey & Erin Spencer (2016). 2016–2017 HR Systems Survey: 19th Annual Edition. *Sierra-Cedar*. Available from Internet: <URL: http://www.sierra-cedar.com/wp-content/uploads/sites/12/2016/10/Sierra-Cedar_2016-2017_HRSystemsSurvey_WhitePaper.pdf>.

Heikkilä, Jukka-Pekka, Chris Brewster & Jaakko Mattila (2014). Micro-Political Conflicts and Institutional Issues During e-HRM Implementation in MNCs: A Vendor's View. In: *Human Resource Management and Technological Challenges*, 1–21. Springer, Cham.

Huang, Liang-Chih, Kuo-Shu Huang, Hsiu-Ping Huang & Bih-Shiaw Jaw (2004). Applying Fuzzy Neural Network in Human Resource Selection System. In: *Fuzzy Information, 2004. Processing NAFIPS'04,* 1, 169–174.

Huang, Min (2015). Forecasting the competitiveness of technological innovations talents using least squares support vector machine. In: *Intelligent Computation Technology and Automation (ICICTA), 2015 8th International Conference,* 429–432.

Huselid, Mark A., Richard W. Beatty & Brian E. Becker (2005). "A players" or "A positions"? The Strategic Logic of Workforce Management. *Harvard Business Review* 83:12, 110–117.

Iles, Paul, David Preece, & Xin Chuai (2010). Talent management as a management fashion in HRD: Towards a research agenda. *Human Resource Development International*, 13:2, 125–145.

Janssen, Marijn, Haiko van der Voort & Agung Wahyudi (2017). Factors influencing big data decision-making quality. *Journal of Business Research* 70, 338–345.

Jantan, Hamidah, Abdul Razak Hamdan & Zulaiha Ali Othman (2009a). Potential Data Mining Classification Techniques for Academic Talent Forecasting. In: *Intelligent Systems Design and Applications, 2009. ISDA'09. Ninth International Conference,* 1173–1178.

Jantan, Hamidah, Abdul Razak Hamdan & Zulaiha Ali Othman (2009b). Classification for Talent Management Using Decision Tree Induction Techniques. In: *Data Mining and Optimization, 2009. DMO'09. 2nd Conference,* 15–20.

Jantan, Hamidah, Abdul Razak Hamdan & Zulaiha Ali Othman (2009c). Knowledge Discovery Techniques for Talent Forecasting in Human Resource Application. *International Journal of Industrial and Manufacturing Engineering* 3:2, 178–186.

Jantan, Hamidah, Abdul Razak Hamdan & Zulaiha Ali Othman (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.

Jensen, Finn V. (2001). *Bayesian networks and decision graphs*. New York: Springer. 258p. ISBN 0-387-95259-4.

Katal, Avita, Mohammad Wazid & R. G. Goudar (2016). Big Data: Issues, Challenges, Tools and Good Practices. In: *Contemporary Computing (IC3), 2013 Sixth International Conference*. 404–409.

King, Kylie G. (2016). Data Analytics in Human Resources: A Case Study and Critical Review. *Human Resource Development Review* 15:4, 487–495.

Kitchin, Rob (2014). *The data revolution: Big data, open data, data infrastructures & their consequences*. London: SAGE. 222p. ISBN 978-1-4462-8748-4.

Le Deist, Françoise Delamare & Jonathan Winterton (2005). What Is Competence? *Human Resource Development International* 8:1, 27–46.

Leiba-O'Sullivan, Sharon (1999). The Distinction between Stable and Dynamic Cross-cultural Competencies: Implication for Expatriate Trainability. *Journal of International Business Studies* 30:4, 709–725.

Lepak, David.P. & Scott A. Snell (2002). Examining the Human Resource Architecture: The Relationships Among Human Capital, Employment, and Human Resource Configurations. *Journal of Management* 28:4, 517–543.

Lin, Hung-Tso (2010). Personnel selection using analytic network process and fuzzy data envelopment analysis approaches. *Computer and Industrial Engineering* 59:4, 937–944.

Lim, Sunghoon, Conrad S. Tucker & Sondar Kumara (2017). An unsupervised machine learning model for discovering latent infectious diseases using social media data. *Journal of Biomedical Informatics* 66, 82–94.

Lismont, Jasmien, Jan Vanthienen, Bart Baesens & Wilfried Lemahieu (2017). Defining analytics maturity indicators: A survey approach. *International Journal of Information Management* 37:3, 114–124.

Lumme-Tuomala, Riitta (2017). *Talent Management in the Humanitarian Aid Context*.

Marler, Janet H. & John W. Boudreau (2017). An evidence-based review of HR Analytics. *The International Journal of Human Resource Management* 28:1, 3–26.

McAfee & Brynjolfsson (2012). Big Data: The Management Revolution. *Harvard Business Review* 90:10, 60–68.

McDonnel, Anthony, Ryan Lamare, Patrick Gunnigle & Jonathan Lavelle (2010). Developing Tomorrow's Leaders – Evidence of Global Talent Management in Multinational Enterprises. *Journal of World Business* 45:2, 150–160.

Minbaeva, Dana B. (2017). Building credible human capital analytics for organizational competitive advantage. *Human Resource Management,* 1–13.

Muhammad, Iqbal & Zhu Yan (2015). Supervised machine learning approaches: A survey. *ICTACT Journal on Soft Computing* 5:3, 946–952.

Nijs, Sanne, Eva Gallardo-Gallardo, Nicky Dries & Luc Sels (2014). A multidisciplinary review into the definition, operationalization, and measurement of talent. *Journal of World Business* 49:2, 180–191.

OrgVue (2018). OrgVue. Available from Internet: <URL: https://www.orgvue.com/>.

Pape, Tom (2016). Prioritising data items for business analytics: Framework and application to human resources. *European Journal of Operational Research* 252:2, 687–698.

Pepermans, Roland, Daniël Vloeberghs & Britt Perkisas (2003). High potential identification policies: an empirical study among Belgian companies. *Journal of Management Development* 22:8, 660–678.

Robinson, Carl, Rachel Fetters, Devon Riester & Anna Bracco (2009). The Paradox of Potential: A Suggestion for Guiding Talent Management Discussion in Organizations. *Industrial and Organizational Psychology* 2:4, 413–415.

Russell, Stuart J. & Peter Norvig (2014). *Artificial intelligence: A modern approach*. 3rd ed. Harlow: Pearson. 1091 p. ISBN 978-1-29202-420-2.

SAP (2018). SAP Successfactors. Available from Internet: <URL: https://www.successfactors.com/en_us.html>.

Saunders, Mark, Philip Lewis & Adrian Thornhill (2012). *Research Methods for Business Students.* 6th ed. Harlow: Pearson. 741 p. ISBN 978–1–273–75075–8.

Saunders, Mark, Philip Lewis & Adrian Thornhill (2016). *Research Methods for Business Students.* 7th ed. Harlow: Pearson. 741 p. ISBN 978-1-292-01662-7.

Scullion, Hugh & David G. Collings (2011). *Global Talent Management*. New York, NY: Routledge. 199 p. ISBN 978–0–415–87171–6.

Shah, Shvetank, Andrew Horne & Jaime Capellá (2012). Good Data Won't Guarantee Good Decisions. *Harvard Business Review* 90:4, 23–25.

Silzer, Rob & Allan H. Church (2009a). The Pearls and Perils of Identifying Potential. *Industrial and Organizational Psychology* 2:4, 377–412.

Silzer, Rob & Allan H. Church (2009b). The Potential for Potential. *Industrial and Organizational Psychology* 2:4, 446–452.

Silzer, Rob & Allan H. Church (2010). Identifying and assessing high-potential talent: Current organizational practice. In Silzer, R.F. & Dowell, B.E. (Eds), Strategy Driven TM: A Leadership Imperative, Jossey Bass, San Francisco, CA.

Sivarajah, Uthayasankar, Muhammad Mustafa Kamal, Zahir Irani & Vishanth Weerakkody (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research* 70, 263–286.

Spreitzer, Grethcen M., Morgan W. McCall & Joan D. Mahoney (1997). Early identification of international executive potential. *Journal of Applied Psychology* 82:1, 6–29.

Stahl, Günter K., Ingmar Björkman, Elaine Farndale, Shad S. Morris, Jaap Paauwe, Philip Stiles, Jonathan Trevor & Patrick Wright (2012). Six Principles of Effective Global Talent Management. *MIT Sloan Management Review* 53:2, 25–32.

Tansley, Carole (2011). What do we mean by the term "talent" in talent management? *Industrial and Commercial Training* 43:5, 266–274.

Tyler, Kathryn (2011). On the Grid. *HR Magazine* 56:8, 67–69.

Ulrich, Dave & James H. Dulebohn (2015). Are we there yet? What's next for HR? *Human Resource Management Review* 25:2, 188–204.

Van den Heuvel, Sjoerd & Tanya Bondarouk (2016). The Rise (and Fall) of HR Analytics: A Study into the Future Applications, Value, Structure, and System Support. *2nd HR Division International Conference*.

Van Knippenberg, Daan, Linus Dahlander, Maetine Haas & Gerry George (2015). Information, Attention, and Decision-Making. *Academy of Management Journal* 58:3, 649–657.

Wei, Sun & Zhao Xiaolin (2011). Evaluation of talent cultivation for colleges and Universities Based on principal component analysis and BP neural network. In: *Information Technology, Computer Engineering and Management Sciences (ICM), 2011 International Conference,* 1, 185–187.

Wiblen, Sharna, Kristine Dery & David Grant (2012). Do you see what I see? The role of technology in talent identification. *Asia Pacific Journal of Human Resources* 50:4, 421–438.

Wilson, Jonathan (2014). *Essentials of Business Research: A guide to Doing Your Research Project.* 2nd ed. Los Angeles, CA: SAGE. 352 p. ISBN 978-1-4462-5733-3.

Workday (2018). Workday. Available from Internet: <URL: https://www.workday.com/en-us/homepage.html>.

Yasodha, S. & P.S. Prakash (2012). Data Mining Classification Technique for Talent Management using SVM. In: *Computing, Electronics and Electrical Technologies (ICCEET), 2012 International Conference,* 959–963.

Ye, Nong (2014). *Data mining: Theories, algorithms, and examples*. Boca Raton, FL: CRC Press. 329p. ISBN 978-1-4398-0838-2.

Yin, Robert K. (2009). *Case Study Research – Design and Methods*. 4th ed. Los Angeles: Sage. 219 p. ISBN 978–1–4129–6099–1.

Zhang, Yong, Guang Li, Wei Guang Li & Ping Ping Xiao (2013). Research on Ability Evaluation Model of Compound Talents of Information Technology Based on Artificial Neural Network. In: *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2013 5th International Conference,* 1, 438–441.

APPENDIX 1.

BACKGROUND INFO
Position:
Tenure:
Experience of working with talent/analytics:

INTERVIEW QUESTIONS
How does your company define talent?

- What is talent like? What words best describe talent?
- Is the definition linked to your company strategy?

How do you identify talent at your company?

- How often and by whom?
- Performance vs. skills, competencies vs. Potential, motivation, aspiration
- What kind of tools are used?
- What kind of measurements are used to identify talent?
- How do you measure potential? How do you measure performance?

What kinds of technologies there are to support talent identification in your company?

- What kind of data is used for these technologies?
  - What challenges are related to data? Data privacy? How have you solved these problems?
- How do you view technology as part of talent identification?
- What do you perceive to be the biggest challenges/benefits when using analytics as part of talent identification?
- How do you see this changing in the coming years?