Implicit Bias when using Data Profiling within Recruitments and Human-Resource Management

Anders Persson

Human-Computer Interaction division, IT-department, Uppsala University, Sweden
anders.persson@it.uu.se

Abstract: Recruiters for organizations and companies use data mining and profiling in large datasets to predict successful candidates. The explicit goal is to get more objectivity in their decisions and selections and to get away from human limitations. The first of three problems that they still might risk to encounter, is to end up on low local peaks of performance maximums, and not be able to adapt to changes. A second risk is to directly or indirectly discriminate candidates on sensitive categories. Finally a third risk is to not give equal opportunities for all individuals, resulting in unwarranted social exclusion. The problems are not new and to some extent they seem similar to what we see with human intuition and implicit biases. It might be an opportunity for implementing change. Two proposed options is to, first, look at scientific methodology and be open to uncertainty, and second, to include diversity in the data mining algorithms. Further research are proposed.

Keywords: Data profiling · Social exclusion · Implicit bias · Recruitment · Profiling · Big Data · Diversity · Discrimination;

1 Introduction

How shall a company looking for a specific person to a specific role, find the most likely candidate to succeed among thousands of applicants? This is a problem facing many recruiters that are either headhunting, or have a job advertisement for popular positions today where it often is 500-1000 applicants to a single job.

To battle this challenge, recruiters within many firms and organizations often use profiling to screen applicants after the likely candidates. Typically it is used as an initial big cut, where some candidates later will be interviewed and evaluated in more detail. If there are no other means at their disposal to make their profiles, recruiters will have to rely on their *intuition.* While specific field-expert's intuitions often can be a very good and quite an exact estimation, it will remain a subjective value assessment, and can be prone to implicit biases for what is deemed "the best candidate". It is often a cause for concern in regards to discriminating factors, for example.

To remedy the limitations of intuition, (big) data mining can be used to make what is claimed more objective profiles for successful candidates [1]. The goal can to some extent be seen as wanting to get away from human limitations, such as subjectivity, and to be more *certain* when making decisions for hiring.

What I will argue for here is that the use of profiling and data mining may get some similar effects as that of implicit biases, as we get from human intuition. They may ultimately have problems with effectiveness, as in delivering on what is promised with the objectivity of the data analysis, and possibly problems with discrimination and giving equal opportunities for applicants. In this paper I will discuss possible ways forward, given that the problems highlighted are true. First, how the problems might be remedied by including *diversity*, and second, to take lessons from the methodology of what is deemed good science, to promote *uncertainty* rather than objectivity and certainty.

2 Data Mining and Profiling

Data mining is a new term, but not necessarily something new. Broadly speaking, it is acquiring useable, meaningful data, out of a larger data set. The term has been invented in recent years, about at the same time as *Big Data* came around, and they remain closely related [2]. Data Profiling is a more specified field within data management, and is: “a specific kind of data analysis used to discover and characterize important features of the data sets. Profiling provides a picture of data structure, content, rules and relationships by applying statistical methodologies to return a set of standard characteristics about data” [3].

In other words, what you get and can use from data mining and profiling is much like data from empirical, statistical research. It might be a risk that the term “data mining” gives the impression of solid, *cold hard facts* dug up from the earth [2]. On the contrary, to get any meaning from statistical data it must be related to a context, and perhaps temporally specified. It also must be interpreted to get any kind of meaning.

2.2 Uniqueness in Data Profiling

An important concept with data profiling is *uniqueness*. One value within a parameter, or set of values in a set of parameters, has maximum uniqueness when it does not have any other sets of values like it. Non-uniqueness will be achieved when two or more sets of values are all the same, and would be the strongest kind of pattern and correlation.

Translated to applicant and their level of academic degree, it would mean that maximum uniqueness is achieved when every candidate has a different level. On the opther hand, if it for example would be a test measure of their programming skill included in the dataset, maximum uniqueness would be when every individual has a different test value. This might become a bit more complicated when you have numbers with decimals, where it might be quite unlikely that two people would get the exact same numbers. But the principle is the same, that minimal uniqueness will be acquired when most individuals are close to each other.

If you would look at the relationship between programming skill as a performance measure, and level of academic degree, you would get *minimal uniqueness* if all those with high programming skills had a *certain level of degree*, and the one’s that scored low a different level of degree. This would make them easy to differentiate, just by looking at their academic level. On the contrary, if programming skill and level of academic degree are *not well correlated*, but varies greatly with good and bad programmer on both levels, *uniqueness is high*.

For a recruiter that wants to be informed how to make selections, you would, then, typically want low uniqueness. It is there where you have more clear differences between applicants and can be more informed when making selections.

2.3 Organizations Using Data Profiling

Using data profiling, or Big Data as it often is just referred to, within Human-Resources (HR) and management is not new. According to a survey in North America in 2014, 73% of companies stated they had, or planned to, invest in Big Data analytic tools [4].

Two examples of bigger organizations using data profiling when hiring are Xerox and Google. I will get to the latter later on in this paper, and will concentrate on the former at the moment.

Xerox is a global business services company with over 130,000 employees. An important element for them to deliver is customer service support. To manage their human-resources within this field they have implemented big data analyses. One explicit performance value they look at is *longevity* within the company, as in how long a person stays within the company [5]. The motivation is, often, that longevity of worker builds organization-specific knowledge over time, and that it is important that this stays within the company. As an analytic tool they have been using Evolv, currently part of a larger platform called Cornerstone OnDemand [6].

They state that they put applicants through a series of tests to predict how they would perform within customer service [5]. This is later followed up in regards to performance measures, such as longevity. Part of the tests are surveys regarding attitudes towards work, such as how long they are willing to travel/commute to work, and how much overtime per week they would be willing to have. Additional information is also gathered from CV's.

Four measures were correlated to longevity, as in staying longer at their job: (1) willingness to work 1-3 hours overtime per week, 15 times more likely, (2) applicants with bachelor's degrees stay 5% longer, and those with technical diplomas stay 26% longer, than those with high school diplomas, (3) those that have had a customer service job where they have had to use empathy, rather than just taking orders, and, finally, (4) living closer to job, or having reliable transportation [5].

It is also claimed that their ideal customer worker is someone that (1) scores good at typing tests, to be better at taking in background documentation on clients, (2) is creative, and (3) uses social media (but not too much). However, it is unclear how this data is correlated to performance [5].

The result is a job profile that human-resource management uses as a predictor of performance within the company. In other words, someone that lives a bit away from the working place, and only has a high school diploma, and does not use social media very much, or at all, ends up with a low predictor and is likely to be cut at an early stage in the hiring process.

Taking it further. It can also be noted that others, but not Xerox, have proposed to use data mining profiles together with expert systems, to help make the selection of applicants [7]. This could be done with more or less automation, either by giving suggestions to a recruiter (a suggestion the recruiter is probably likely to follow), or the system itself making initial selections.

No example has been found where recruiters use extracted content data from social media, to include in a job profile predictor. However, a study in Sweden showed that at least half of interviewed recruiters in big and middle-sized organizations did scan applicant social media profiles themselves, at some point before hiring [8]. In the survey they stated they looked for risk-behaviors, if they shared the organization's values, and things that might damage a company's reputation. It would not be a difficult step to include data from public, obtainable, social media profile of applicants, and make a text-feature analysis to capture potential risk behavior.

3 Potential Problems

Using profiles to cut a big portion of applicants might be seen as an effective way of siphoning out the less likely performing candidates, and mere "fair and square" competition between applicants. On the face of it companies should be able to perform better, with more people at the right place, and companies like Xerox report results going upwards after using data profiling for hiring.

I will present some potential problems in this next section, regarding effectiveness long-term, and some regarding ethics, to be discussed further on in the paper.

3.1 Local Maximums and Stagnation

An ongoing research question within product development concerns innovation and success in different economic contexts. To shed light on this matter, some have applied an evolutionary perspective, and something called a fitness landscape [9]. Within this concept there is a difference between local and global maximums. An organism, with certain attributes at its disposal (a certain genotype) adapting to a certain environment, is theorized to sooner or later end up on a local peak (performance) optimization, for that environment.

Related to humans and organizations, this could mean that given a certain type of workforce you will end up at a certain level of performance. Within your own workforce you should also be able to see different levels of peaks, with a large enough dataset. This could probably be done in larger organizations with (big) data analyses. You would then typically want to identify where the peak performances are, and spread *those type abilities* those people have to other positions in the same occupation.

However, when changing to a different environment, the same set of optimized abilities may not be as effective or competitive compared to other sets of abilities. All data that is mined, and all data of empirical research alike, is historical data. They do not really say anything about the future. Given that the global market, influenced by people and societies all around the globe, can have quickly changed conditions for success, this could be problematic. This is a problem for any use of profiling, of course, not just data mined.

I will discuss a potential solution with re-profiling below.

3.2 Discrimination

The most ethically relevant problem is probably that of potential discrimination in the selection process. Discrimination is controlled by law, in most countries, and is usually defined as: “the treatment of a person or particular group of people differently, in a way that is worse than the way people are usually treated” [10]. Expressed in a different way, it is judging people based on things as social group categories, rather than individual merit. Most commonly this is an issue with features such as: (1) gender, (2) ethnical background, and (3) sexual orientation.

It is important to note that discrimination can be direct and based on discriminatory categories, such as those mentioned. It can also be indirect; a selection based on non-discriminatory categories, which, however, are strongly correlated to the discriminatory one’s [11].

There are, for example, reports of companies scanning employee’s health records, to predict those likely to get on sick-leave, or those that might be pregnant [12].

3.3 Loss of Opportunities

Finally, a potential problem with the use of data profiling to predict applicants performance could be a *loss of opportunity* for the outliers in the features included. That is, the individuals that differ from the norm, and predicted most likely high performer, but still is just as likely to perform at the same level. This is likely, given the premise that measures for the applicants have normal distributions.

This is to some extent related to uniqueness in data profiling. Given that if the more informative measures related to performance will have less uniqueness, they will also be less forgiving towards outliers. Those outliers from the norm could be seen as loosing opportunities, and this loss could be an ethical problem of *social exclusion*.

Another concern is also that things outside the work-sphere are made relevant when judging work performance. Xerox was, for example, looking for people with mid-range social media use, for an occupation as customer service. Using social media is not, at least, directly related to the task they will do, but apparently they might related indirectly. It could possible exclude applicants on the sole difference of not using social media, such as Facebook. Given that people are adapting to the demands on job markets they want to get into, it might mean that people conform their private life, as well.

3.4 Reprofiling for Quality of Data

One approach to remedy some of the problems just mentioned might be simply to seek better quality of the data. It could be claimed that the problems of local maximums could be avoided if one continuously updated the profiles, to what kind of workers one want at the specific time and date. Even more so, if the quality of the (meaningful) data extracted from the big data set is good enough, and detailed enough, it should be able to pick out those individual outliers that still are high performing.

It does not seem to be an unreasonable claim, but it might be problematic in practice, given that repeated use might already have weeded out the outliers. I will relate to this process in the next chapters, about intuition, implicit bias and recruitment for female executives.

Data quality is first and foremost an issue when it comes to, so to speak, living data sets that change over time (such as what correlates to performance within a certain organization on renewed global markets). To remedy this, most data sets can benefit of what is called *reprofiling;* a new analysis of the dataset to see if statistical patterns and relations has changed [3].

However, a risk, as far as I can analyze it, is that instead of including outliers it might rather exclude them further. Let's say we mine out a profile P1, with features such as those of Xerox are looking for in customer service personal. You then choose people from profile P1 to be added to the workforce for the specific occupation, and then re-profile with the new selection of individuals added, to a profile P2.

If nothing significant has changed in regards to what measures are correlated to high performance between P1 and P2, the pattern will have statistically have become "stronger"; as in, with less deviations from means, and with lower p-values. This will most likely mean that P2 would be even less forgiving to outliers, unless your data analysis is detailed enough. Big organizations like Xerox, or Google as I will come later, with abundant resources at their hand might be able to make it detailed enough, but would probably be difficult for others.

Revisiting Local Maximums. So, what does this then mean for the problem of ending up on low local peaks of performance with changing market-environment. You would after some time have few outliers to rely on if profile P1, or P2, has changed the workforce according to it. With fewer outliers to rely on, it will be more difficult to find the pattern for the new local peaks of performance, in the new environment. The workers who would be better suited for this new environment might simply not he left in the company.

On the face of it, then, reprofiling might not be enough to solve the problem of local maximums. I would like to note two premises for this effect to be relevant: (1) profiling and reprofiling is made on a closed group, as in not including a larger part of a specific market environment, (2) that using profile P1 actually makes the pattern P2 have lower significances and stronger correlations to performance measures, and (3) that the analysis is not matching outlier results in the data profiling to new applicants. Otherwise the theorized effect would not be there.

3.5 Human Intuition and our Predictive Mind

The problem with profiling and reprofiling might not be too different from what we initially wanted to get away from; the human, intuitive way of reasoning and categorizing, in a limited and subjective fashion. More specifically it could be related to what is often called *implicit bias*, a term often used in relation to discrimination and how our brain seems to function when unconsciously categorizing.

An explanatory theory within neuroscience that has gained momentum lately is one where the brain essentially is seen as a hypothesis-testing mechanism, and that we do a kind of statistical analysis within our neurons [13]. The main goal, so to speak, for the brain is to be able to predict what will happen, before it happens. Within the connections between neurons we have both a forward progressive movement; taking in input from nerve-endings, like our eyes, to be processed in succession in our neural cortex. But we also have a backward progression, signals constantly sent backwards, from higher-cognitive areas of the brain, towards lower. This has been explained as a kind of feedback-loop, but lately it has been given more emphasis and rather explained it the other way around.

Rather than starting blank and take in input to create a model of our surroundings, we start with a working model that is projected backwards within our connective network. The signal input then functions as an error-correction function, to adjust our working model of the world around us where it is needed in the moment, but first and foremost we already have a model set. Instead of giving feedback, the brain, first and foremost, gives errors to an existing model of the world.

Studies also show that the kinds of categorizations we do to learn about the world are very much like statistical analyses. We more or less make a *mirrored casual structure of the world* in our brain, from what we have experienced [13].

This can be related to something called confirmation bias in perception; we have a tendency to direct our attention, unconsciously, towards that which confirms our prior beliefs. Related to the theory of our predictive mind, in a backward loop we project a working model of our surroundings, and our attention is drawn towards that which confirms this model to be true. In a sense then, our prior beliefs, and prior categorizations we have made from experience, determines what kind of info we will look for when facing a new situation; that which confirm what we already know.

In some sense this can be related to profiling and reprofiling. After making a categorization, a profile, it will direct attention only to the feature you know they are related to. Even if you try to reassess you prior knowledge, as in if you reprofile, you will end up looking at the same features, and probably reinforce your prior belief. A risk with human intuition is that prior beliefs becomes dominating, and determines also your future beliefs.

Implicit bias. What seems to happen with the described function of our predictive mind is essentially what is happening with implicit biases; an attempt for our brain to predict future outcomes, based on previous experience and working models. For example, several studies have examining how different individuals judge a written application. The only differences in the studies have been the name of the applicant; one native and one foreign [14]. In US as well as Sweden, respectively, the applicant with the native name on the application is much more likely to be called to an interview. The one with a foreign name is, with the same merits, less likely to be called.

There is a similar effect of women being hired for management positions. A very clear example of the effect of how implicit bias could work is the following:

> “There is good news and bad news about actual gender-related managerial differences. The good news is that some do exist. The bad news is that they are overused as the basis for sexual stereotyping.” [15]

This might be a bit of a controversial find, that it seemed to be a difference between female and male performance in abilities related to management positions. However, they also note that they are overused when recruiting. This is very much what can be seen happening when over-categorizing. While there might be a difference in performance, the difference is not black and white or clear-cut.

This becomes more clear with an example of gender difference in math-skill. SAT-results in US from 2013 confirms that male, for one or the another reason, performed better; an average of 531 (SD = 121) for men, versus 499 (SD = 114) for women [16]. However, given the standard deviations, 38% of female participants outscored 50% of the male participants.

So, if we have one male and one female in front of you, and have task of selecting the one we deem most likely be best in math, the difference in score is not very informative. An almost 40% chance to be wrong if choosing the one associated to the group with highest mean, are not very good odds. Still, most people’s intuition will be much more certain, and a gender based stereotyping and over-categorization will remain.

Revisiting the customer service profile. Same thing can perhaps be said about some of the features on the profile from Xerox and their customer service personal. While it may be more likely that applicants with a bachelor’s degree is, significantly, likely to stay 5% longer than those with high school diplomas, there might be many with the high school diplomas that will stay longer than the them.

Once again relating to the prediction-error theory of our mind, data mining and profiling can be seen as an attempt to predict the future, based on historical experience (data). To continue on that analogy, it could be seen as if (big) data profiling can have implicit biases, or reinforce the idea that humans through perception of the world are (imperfect) statistical machines.

4 Doing it differently

What I have argued for so far are potential risks for using data profiling when hiring. As I also have tried to show is that they are not necessarily any new risks, but can be seen in any use of profiles. It even seems to relate to the way our human intuition categorizes from experience, to try to predict future outcomes.

Next I will try to look at how it could be done differently, mainly looking at two aspects; promoting a bit of uncertainty instead of a quest for objectivity, and to including diversity in the mix.

4.1 Promoting Uncertainty

As was touch on in the beginning, recruiters and management can be seen as seeking very much support for their decisions, hoping to be more *certain* when hiring someone [1, 7]. Perhaps what is needed is a bit of the opposite, and to look at methodology of how *good science* is performed.

Key aspects are to make hypothesizes, so called null hypotheses; you seek to disprove a theory, rather than proving it. Rather than looking to confirm your theory at hand (your working model, if you like), you try to deny it from being true. So, rather than certainty there is a promotion of *uncertainty*.

Granted, to some extent data profiling can be seen as trying to achieve exactly what is proposed. Recruiters do not want to fool themselves, and organizations do not want to underperform. The motivation to go for Big Data is very much to disprove previously held theories of what is most effective and what is not, and if necessary, revamp their whole approach.

Nonetheless, what this could mean is that you as a recruiter should activity try to disprove your own hypotheses. Not just once, but continuously if you want to avoid getting stuck on low peaks of performance when the environment around you change. As will be seen with an example further down, you could also strive for looking for alternative scenarios, and try to predict future conditions for high performance.

4.2 Including Diversity

Another way of solving getting stuck at low local maximums could be to make sure your workforce has enough diversity to not stagnate when conditions change.

A warranted question is if that would be enough to remain innovative and adaptive. Some where they have seen a correlation between occupational diversity and likelihood of innovating research, seems to suggest that it is [17]. Another study also found that openness to cultural diversity was correlated to likelihood for companies continued economic performance, which was interpreted as being able to renew itself and to innovate [18].

To include diversity could, of course, also solve the problem of discrimination. It might depend on what kind of categories, and how, you make them diverse. For example, you could perhaps at least try to conform to actual differences between genders. For example, let’s say we have a specific task to find someone solely based on math skill. Related to the example of gender differences in SAT-math score in section 3.4, to implement diversity could be to, at least, result in similar distributions as with the test scores. In other words, you would wanna see about 38% women and 62% males to be chosen, instead of over-generalizing to explicitly look for male applicants. Any man will not be the most likely best mathematician.

Same thing, then, could be applied to the difference between academic degree and customer service workers within Xerox (see section 2.3). Without having specific numbers or distributions available, but given that also they follow normal distributions as most data about humans will, you would want to, at least, see similar spread and diversity in applicants you accept. Those with bachelor's degree will be more in numbers, but not exclusively.

This could also be a possible way to open up for equal opportunities for people that are just as likely to perform well within a certain occupation, that otherwise would not have gotten the chance. In a very direct way the purpose can be said to be to not limit uniqueness too much in your data as well as with individuals, when including diversity.

Possible critic. If this is the case, that diversity is wanted, you could ask yourself why profiling at all. Why data profiling at all, if it is good to have a diverse workforce instead of one with certain, highlighted, high performing features?

To some extent it might be question of balancing between long-term and short-term goals. On one hand, the data probably don't lie, and if a company *adopts* the profile they get from current data, the company would probably perform well in the near future. On the other hand, it might become difficult to *adapt* to changing conditions, and will not perform as well long-term.

4.3 Google Showing a Way

Perhaps an organization that to some extent implements some of these principles can be found within Google's Human-resource management procedures. They have in recent decade implemented data profiling and big data analyses to create predictive models for hiring, as well as seeking to reorganize employees where they would fit best.

They state that they actively try to seek out root causes where there is "weak diversity", as with under-representations of female engineers [19]. They also state that they do predictive modeling with "what if" scenarios, to try to forecast what might happen further on [19]. Exactly what is meant with this is unclear with the present literature available, and is due to further research. To some extent they could, perhaps, tackle some of the issue of getting stuck on low peak performances of local maximums, and perhaps also tackle concerns for discrimination and equal opportunities.

They also seem to have longevity in mind, as Xerox has (see section 2.3), with their own *retention algorithm*; "a mathematical algorithm to proactively and successfully predict which employees are most likely to become a retention problem" [19]. Once again, exactly how this is to be understood is unclear at the moment of writing, and what it means that an *employee becoming a problem* and what proactive measures that are done, is also unclear. The way it is phrased, at least, might raise some cautionary eye-brows of ethical concern, but that is beyond the scope of this paper.

5 Conclusions

What this paper tries to do is to present an initial investigation of how recruiters use Big Data and data profiling when hiring. The use of profiles from mining algorithms often has the explicit purpose of getting *objective* values for what increases performance for a company or organization. Some potential problems that they risk to run in to have been highlighted here, the first being to end up on low peaks of local maximums and miss out on achieving long term performance and global maximums. Second is a concern for discrimination when using data mining algorithm, third and final a related concern for a loss of equal opportunities.

The problems discussed here are not new. They are proposed to even be related to problem you can see within our own human cognition and intuition. There we have an unconscious tendency to over-categorize, and actively seek to confirm our prior beliefs. In some sense the problems in our cognitive mind might highlight some ways going forward with the concern for data profiling and recruitment.

A potential remedy discussed is to look for uncertainty, rather than certainty and objectivity. Rather than cold, hard, unquestionable facts, what you could try to emulate is methodology for good science. You could question the validity of a profile you get when data mining, and instead look for alternatives, to dig deeper and to try to disprove its legitimacy with varied environments and contexts. It will remain a difficult task to try to predict what will spell *success* in the future. It is the same reason why science limits itself to investigate what has happened, and try to say very little if anything at all, about the future. Nonetheless, everything can be more or less predictable, also in science.

Another remedy could be to keep a diverse workforce for an organization, to be able to continually innovate in a changing, global, competitive market. How this is to be done and balanced would have to be investigated further, but perhaps it is possible to implement diversity algorithm directly into the data profiling algorithm.

This way you might be able to make sure to avoid the risk of ending up discriminating based on sensitive categories, as well as to some extent keep equal opportunities open to all that deserve to have it. Summarized, it might be good to keep a keen eye on individuality and uniqueness, in data as well as in with individuals.

5.1 Further Investigations

This paper is mostly a hypothesized and theorized conclusion, and for further investigations in this problem space it would be needed more empirical data of how organizations, human-resource managers, and recruiters use data profiling. As no detailed data, or studies, has been able to been reviewed from neither Xerox, nor Google, further research would first and foremost be with them.

Some of the things argued for in this paper could also be simulated and tested. For example, the effect argued for with reprofiling; that statistical patterns would get stronger, in the sense of resulting in less deviation from means for individual features, and get lower p-values.

It is also a bit unclear how diversity could go hand in hand with profiling with specific and limited features. Both can seem to be contradicting, as the former is looking for more uniqueness and spread on data, and the latter probably want less uniqueness to be informative in applicant selections.

Acknowledgements. I would like thank my associate supervisor Mikael Laaksoharju, here at Uppsala University, for extensive feedback and contributions; both with providing examples to use, first and foremost about local maximums, as well as structure and presentation of the paper. His keen eye would probably have had something to say about the long sentence he was just mentioned in.

References

1. Chien, C-F., Chen, L-F.: Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. Expert Systems with Applications, volume 34, pp. 280-290 (2008)
2. Portmess, L., Tower, S.: Data barns, ambient intelligence and cloud computing: the tacit epistemology and linguistic representation of Big Data. Ethics and Information Technology, Volume 17, p. 1–9. DOI: 10.1007/s10676-014-9357-2 (2015)
3. Sebastian-Coleman, L.: Measuring Data Quality for Ongoing Improvement: A Data Quality Assessment Framework. Morgan Kaufmann Publishers, Burlington, Mass. (2013)
4. Gartner, Inc. Technology research, viewed 2016-07-22, http://www.gartner.com/newsroom/id/2848718
5. Quartz, viewed 2016-07-22, http://qz.com/146000/big-data-can-take-the-guesswork-out-of-the-hiring-process-2/
6. Cornerstone OnDemand, viewed 2016-07-22, https://www.cornerstoneondemand.com/evolv
7. Mehrabad, M. S., Brojeny, M. F.: The development of an expert system for effective selection and appointment of the jobs applicants in human resource management. Computers & Industrial Engineering, Volume 53, pp. 306-312 (2007)
8. Backman, C., Hedeneus, A.: Will your Facebook profile get you hired? Employers use of information seeking online during the recruitment process. The 6th Biannual Surveillance and Society Conference, Barcelona, Spain (2014)

9. Kwasnicki, W., Kwasnicka, H.: Market, innovation, competition, An evolutionary model of industrial dynamics. Journal of Economic Behavior and Oranization, vol. 19, pp. 343-368 (1992)
10. Cambridge Dictionaries Online, retrieved 2016-07-22, http://dictionary.cambridge.org/dictionary/english/discrimination
11. Hajian, S., Domingo-Ferrer, J.: A Methodology for Direct and Indirect Discrimination Prevention in Data Mining. IEEE Transactions on Knowledge and Data Engineering, volume 25, issue 7, pp. 1445-1459 (2012)
12. Dailymail, The Mail on Sunday & Metro Media Group, retrieved 2016-07-22, http://www.dailymail.co.uk/news/article-3456691/How-companies-secretly-tracking-employees-health-private-lives-big-data-save-money-lead-big-problems.html
13. Hohwy, J.: The Predictive Mind. Oxford University Press, Oxford (2013)
14. Jost, J. T., Rudman, L.A., Blair, I.V., Carney, D.R., Dasgupta, N., Glaser, J., Hardin, C.D.: The existence of implicit bias is beyond reasonable doubt: A refutation of ideological and methodological objections and executive summary of ten studies that no manager should ignore. Research in Organizational Behavior. Volume 29, pp. 39-69 (2009)
15. Sexton, D.L., Bowman-Upton, N.: Female and male entrepreneurs: Psychological characteristics and their role in gender-related discrimination. Journal of Business Venturing, volume 5,iIssue 1, pp. 29-36 (1990)
16. Psychology Today, Sussex Publishers, retrieved 2016-07-22, https://www.psychologytoday.com/blog/good-thinking/201403/why-the-gender-difference-sat-math-doesnt-matter
17. Söllner, R.: Human Capital Diversity and Product Innovation: A Micro-Level Analysis. Jena Economic Research Papers, 027 (2010)
18. Ostergaard, C.R., Timmermans, B., Kristinsson, K.: Does a different view create something new? The effect of employee diversity on innovation. Research Policy, volume 40, pp. 500-5009 (2011)
19. Eremedia, retrieved 2016-07-22, http://www.eremedia.com/ere/how-google-became-the-3-most-valuable-firm-by-using-people-analytics-to-reinvent-hr/